

# Non-parametric Estimation of the Death Rate in Branching Diffusions

R. HÖPFNER

*Johannes Gutenberg Universität Mainz*

M. HOFFMANN

*LPMA and Université Paris VII*

E. LÖCHERBACH

*Université Paris XII*

**ABSTRACT.** We consider finite systems of diffusing particles in  $\mathbb{R}$  with branching and immigration. Branching of particles occurs at position dependent rate. Under ergodicity assumptions, we estimate the position-dependent branching rate based on the observation of the particle process over a time interval  $[0, t]$ . Asymptotics are taken as  $t \rightarrow \infty$ . We introduce a kernel-type procedure and discuss its asymptotic properties with the help of the local time for the particle configuration. We compute the minimax rate of convergence in squared-error loss over a range of Hölder classes and show that our estimator is asymptotically optimal.

*Key words:* branching diffusions, kernel estimation, minimax estimation

## 1. Introduction

Over the last decades there has been a growing demand for developing methods of statistical inference in complex models, motivated by applications in biology, in particular in genetics. The theory of (classical) branching processes has long been used to model evolutionary problems and is well understood both from a probabilistic and a statistical point of view (see Athreya & Ney (1972) and Guttorp (1991), for a fairly recent account on statistical inference for classical branching processes). However, simple branching processes do not permit modelling of spatial properties, e.g. if individuals or particles are subject to motion, or, from a biological point of view, if a selection process varies geographically. Likewise, it may be relevant to consider interactions between particles.

Models of type “stepping stone” allowing discrete time and space (i.e. introducing a discrete motion or a discrete distance between mutations in genes) were considered by, Kimura & Weiss (1964), Malécot (1967) and several other authors. First generalizations to either time or space (or both) being continuous, the simplest continuous migration law being Brownian motion in  $\mathbb{R}^d$ , were considered by Fleming & Su (1974) and Nagylaki (1978). These generalizations lead us to the model of spatially branching diffusions. As Stanley Sawyer explains in Sawyer (1976): “Probably the best application of branching diffusions in genetics is describing dispersion, mutation and geographical selection of the descendent of a new gene in a uniform high-density population, or any population of rare mutant genes.” See also Ewens (1969) or Crow & Kimura (1970). Another (mathematically equivalent) application of branching diffusions is a model of epidemics in a large population with geographical structure (Neyman & Scott, 1964).

Spatially branching diffusions are fairly well understood from a probabilistic point of view, see e.g. Etheridge (1993), Wakolbinger (1995) and the references therein. However, statistical inference has not reached a unified form yet and only several particular cases of the general model of branching diffusions have been covered. To mention but a few: Adke (1983) studies

the MLE theory in the specific case of a branching Brownian motion with no deaths, constant branching rate and reproduction mechanism which always creates  $k = 2$  offspring, for various sampling schemes of observation. Kulperger (1986) performs least-squares parametric estimation in branching diffusions with constant branching rate, constant diffusion coefficient and zero drift, but for a general offspring mechanism and including immigration. His model is the closest form to the general study we propose in this paper. More recently, a characterization of the parametric structure (the LAN or LAMN property) of general branching diffusion models has been given in Löcherbach (see Löcherbach, 1999a, b, 2002) in a framework of particle processes where coexisting particles move, branch and reproduce dependently.

### 1.1. Branching diffusions with immigration

In this paper, we study a specific estimation problem for a wide class of branching diffusions with immigration. The branching diffusion model with immigration that we have in mind can be described as follows. A finite system of particles lives in a one-dimensional space and is subject to the following dynamics:

1. Particles move, independently of each other, on diffusion paths

$$d\xi_t = b(\xi_t)dt + \sigma(\xi_t)dW_t. \quad (1)$$

2. They branch at position-dependent rate  $\kappa(\cdot)$ , according to a reproduction law  $p_k(\cdot)$ ,  $k = 0, 2, 3, \dots$ . More specifically:
  - (i) A particle sitting in  $x$  at time  $t$  will die in a small time interval  $]t, t + h]$  with probability  $\kappa(x)h + o(h)$ .
  - (ii) At its death time the particle is replaced at its death position  $x$  by  $k$  offspring particles with probability  $p_k(x)$ ,  $k = 0, 2, 3, \dots$
  - (iii) The offspring particles then start independent motions according to (1) and branch again after position-dependent killings.
3. The population is alimented by an immigration process. Immigrants enter the population in random time and space positions governed by a Poisson random measure.

### 1.2. Objectives

After having explored the parametric structure of branching diffusions in Löcherbach (1999a, b, 2002), the next logical step in the statistical investigation of branching diffusion models is non-parametric estimation: the complex nature of the non-parametric approach, i.e. allowing the unknown parameters to be infinite dimensional, requires a good understanding of the finite dimensional—the parametric—case. Besides, in a real case study, a practitioner would presumably first use a non-parametric approach in order to validate the model that has to fit the data. This becomes even more legitimate for the biological applications we have in mind: up to now, scientists do not seem to have a clear view of the possible underlying parametric structures in branching diffusions (see for instance Holmes & Lewis, 1994, for a survey on models used in ecology).

The goal of this paper is to explore the statistical complexity of the model described above in a non-parametric setting. We aim at giving precise statistical characteristics (such as optimal rates of convergence, estimation procedures) in order to compare branching diffusions to benchmark non-parametric models like density estimation or non-linear regression.

To be more specific: given the continuous observation of the system of particles living according to a branching diffusion with immigration dynamic over a time interval  $[0, t]$  (with asymptotics being taken as  $t \rightarrow \infty$ ), we infer on the different parameters of the model which are: the diffusion coefficient  $b$ , the branching or killing rate  $\kappa$ , the reproduction law  $p$  and the Poisson immigration process. (In a continuous time setting, the diffusion coefficient  $\sigma$  is known (it can be identified by the quadratic variation of the diffusion paths  $\xi$  which is an observable quantity).) A typical problem is the estimation of the unknown branching rate  $\kappa(\cdot)$ —which is the goal of this paper. Once this goal is achieved, estimation of a spatially dependent drift or a spatially dependent reproduction law via kernel estimates can presumably be done *mutatis mutandis* using the same techniques as developed in this paper. Moreover, since the immigration process is a Poisson process, independent of the particle's motion, the problem of estimating the immigration rate or the immigration measure is completely covered by known results on estimation problems for point processes (see for instance Kutoyants, 1998, and the references therein).

Therefore we restrict our attention to the estimation of the unknown branching rate  $\kappa(\cdot)$ .

### 1.3. Organization of the paper

In section 2, we present the statistical setting and our results: we measure smoothness of the branching rate by considering Hölder classes of possible shapes for the branching rate and suppose that the branching rate has smoothness of order  $\beta$  in a Hölder sense. For point-wise estimation, we obtain convergence of the error of the Nadaraya–Watson estimator at rate  $\sqrt{th}$  to a centred normal limit law as soon as the bandwidth  $h$  of the kernel estimate is  $o(t^{-1/(2\beta+1)})$  (theorem 1.b). Allowing a tight sequence of bias terms, we can relax the bandwidth condition to  $h = O(t^{-1/(2\beta+1)})$ , and thus the rate of convergence of our estimate is the classical  $t^{-\beta/(2\beta+1)}$ . We also show that the rate  $t^{-\beta/(2\beta+1)}$  is asymptotically optimal for the minimax risk in an integrated  $L^2$ -error over compact intervals (theorem 2). In section 3, we develop new probabilistic tools which are necessary for proving the upper bound: this includes a monotonicity result for the invariant measures with respect to the branching rate  $\kappa$  (theorems 3 and 4). We also introduce a local time for the particle process (section 3.2) and obtain its asymptotics via a Tanaka formula. Section 3 can be read independently from section 2 and has some interest for its own. Section 4 is devoted to the proofs.

## 2. Statistical setting and main results

### 2.1. Branching diffusions with immigration

We consider a stochastic process  $\varphi = (\varphi_t)_{t \geq 0}$  taking values in the space  $S = \cup_{l \geq 0} \mathbb{R}^l$  of finite (ordered) configurations of particles in  $\mathbb{R}$ , whose paths are piecewise continuous  $\mathbb{R}^l$ -valued functions, for varying  $l \in \mathbb{N}_0$ . Our assumptions are the following:

#### Assumption 1

During the random lifetime of an  $l$ -particle configuration, the particles belonging to it travel in space according to a solution  $\xi = (\xi_t^i)_{1 \leq i \leq l}$  of the system of SDE

$$d\xi_t^i = b(\xi_t^i)dt + \sigma(\xi_t^i)dW_t^i, \quad \xi_0^i = x^i \in \mathbb{R}, \quad 1 \leq i \leq l$$

with independent one-dimensional Brownian motions  $W^1, \dots, W^l$ , where  $b$  and  $\sigma$  are globally Lipschitz on  $\mathbb{R}$ , and where  $\sigma$  is strictly positive.

**Assumption 2**

1. *Independently of each other, particles are killed at position dependent rate*

$$\kappa : \mathbb{R} \rightarrow (0, \infty).$$

$\kappa$  is continuous, bounded, and bounded away from 0.

2. *The particles leave a random number of descendants at their death position, according to a family of position dependent reproduction laws*

$$p(\cdot) = (p_k(\cdot))_{k \in \mathbb{N}_0 \setminus \{1\}}, \quad p: \mathbb{R} \rightarrow \mathcal{M}^1(\mathbb{N}_0 \setminus \{1\}).$$

*These descendants then start independent motions according to assumption 1.*

3. *The functions  $\varrho(x) = \sum_{k \neq 1} k p_k(x)$  and  $\tau(x) = \sum_{k \neq 1} k(k-1)p_k(x)$  are finite and continuous on  $\mathbb{R}$ , and  $\tau$  is bounded.*

4.  *$p$  is continuous in its space variable and admits some fixed space-independent “subcritical” law  $\bar{p} \in \mathcal{M}^1(\mathbb{N}_0 \setminus \{1\})$  such that  $\sum_{k \neq 1} k \bar{p}_k < 1$  as an upper bound on  $\mathbb{R}$  in the sense of convolution of probability measures:*

$$\forall x \in \mathbb{R}, \text{ there exists } \hat{p}(x) \in \mathcal{M}^1(\mathbb{N}_0 \setminus \{1\}) \text{ such that } p(x) \star \hat{p}(x) = \bar{p}.$$

**Assumption 3**

*New particles (one immigrant per immigration time) immigrate at constant rate  $c > 0$  and choose their position in space according to a probability law  $\pi'$  on  $\mathbb{R}$ . We assume that  $\pi := c\pi'$  admits a continuous Lebesgue density denoted by  $r$ .*

As a special case of the construction given in Löcherbach (1999a, b), processes meeting assumptions 1–3 exist: there is a closed subset  $\Omega$  of the Skorokhod space  $D(\mathbb{R}^+, S)$  such that all trajectories  $\omega \in \Omega$  have only jumps either of immigration type or of branching type. (Note that  $\omega \in \Omega$  has only “big” jumps changing the length of the configuration, in particular, there is no accumulation of jumps in finite time.) Write  $\varphi, \varphi_t(\omega) = \omega(t), t \geq 0, \omega \in \Omega$ , for the canonical process on  $\Omega$ , endowed with its Borel  $\sigma$ -field  $\mathcal{A}$  and with the canonical filtration  $\mathbb{F} = (\mathbb{F}_t)_{t \geq 0}$  where  $\mathbb{F}_t = \cap_{r > t} \sigma\{\varphi_s : 0 \leq s \leq r\}$ . For every set of “parameters”  $b, \sigma, \kappa, p, \pi$  as in assumptions 1–3 and for every  $x \in S$ , there is a unique probability law  $Q_x^{b, \sigma, \kappa, p, \pi}$  on  $(\Omega, \mathcal{A}, \mathbb{F})$  such that under  $Q_x^{b, \sigma, \kappa, p, \pi}$ , the canonical  $\varphi$  is strongly Markov with initial configuration  $x$  a.s., with a sequence of successive jump times  $(T_n)_n$  increasing to infinity, such that all possible arrangements of a new configuration after jump times  $T_n, n \geq 0$ , are equally probable. We call this process  $\varphi$  on  $(\Omega, \mathcal{A}, \mathbb{F}, Q_x^{b, \sigma, \kappa, p, \pi})$  a branching diffusion with immigration (BD).

Due to the strong assumptions on  $\kappa$  and  $p$  in assumption 2, the BD processes considered here are ergodic and admit the void configuration  $\Delta$  as a recurrent point (Löcherbach, 1999a, th. 6.11.d). Writing  $R_1, R_2, \dots$  for the successive entry times to the void configuration,  $R_0 \equiv 0, R := R_1$ , the unique invariant probability for  $\varphi$  under  $Q := Q_x^{b, \sigma, \kappa, p, \pi}$  is  $m := m^{b, \sigma, \kappa, p, \pi}$

$$m(F) = \frac{1}{E_\Delta(R)} E_\Delta \left( \int_0^R 1_F(\varphi_s) ds \right), \quad F \in \mathcal{S} \tag{2}$$

where  $\mathcal{S}$  denotes the Borel  $\sigma$ -field on the configuration space  $S$ .

Interpreting a configuration  $x \in S$  as a measure on  $(\mathbb{R}, \mathcal{B})$ , with  $\mathcal{B}$  the Borel  $\sigma$ -field on  $\mathbb{R}$ , we write  $x(B)$  for the number of points of  $x$  belonging to  $B \in \mathcal{B}$ . For  $f$  non-negative and measurable,  $x(f)$  is the integral  $\sum_{i=1}^{l(x)} f(x^i)$  if  $x$  is a configuration  $(x^1, \dots, x^{l(x)})$  with  $l(x) > 0$ . For the void configuration  $\Delta$ , we put  $\Delta(f) = 0$ . We define a measure

$$\bar{m}(B) = \frac{1}{E_{\Delta}(R)} E_{\Delta} \left( \int_0^R \varphi_s(B) ds \right), \quad B \in \mathcal{B}. \tag{3}$$

This is—up to norming—the expected amount of time visited in  $B$  by all particles whose life span is contained in one fixed life cycle of  $\varphi$ . We suppose that  $\bar{m}$  is a finite measure, see also assumption 4 below; we call it the invariant measure on  $(\mathbb{R}, \mathcal{B})$ .

2.2. Construction of the estimator

In order to construct the estimator, we need some notation: we decompose the sequence of jump times  $(T_n)_n$  of  $\varphi$  into a sequence of immigration times  $(T_n^I)_n$  and a sequence of branching times  $(T_n^B)_n$ :

$$T_0^I \equiv 0, \quad T_{n+1}^I := \inf\{T_m > T_n^I : \varphi_{T_m} = \varphi_{T_m^-} + \varepsilon_y \text{ for some } y \text{ s.t. } \varphi_{T_m^-}(\{y\}) = 0\} \quad \text{for } n \geq 0,$$

where  $\varepsilon_y$  is the Dirac measure at  $y \in \mathbb{R}$ , and

$$T_0^B \equiv 0, \quad T_{n+1}^B := \inf\{T_m > T_n^B : T_m \neq T_k^I \text{ for all } k\} \quad \text{for } n \geq 0. \tag{4}$$

Then immigration positions

$$\zeta_n^I := \text{supp}(\varphi_{T_n^I} - \varphi_{(T_n^I)^-}), \quad n \geq 1$$

are well defined, and  $\zeta_n^I$  is  $\mathcal{F}_{T_n^I}$ -measurable. The branching positions are

$$\zeta_n^B := \text{supp} \left( \frac{1}{k-1} (\varphi_{T_n^B} - \varphi_{(T_n^B)^-}) \right) \quad \text{if } l(\varphi_{T_n^B}) = l(\varphi_{(T_n^B)^-}) + k - 1, \tag{5}$$

$n \geq 1, k \neq 1$ , with  $\zeta_n^B$   $\mathcal{F}_{T_n^B}$ -measurable. We get  $\mathbb{F}$ -multivariate point processes of immigration resp. branching times/positions:

$$\begin{aligned} \mu^I(ds, dy) &= \sum_{j \geq 1} 1_{\{T_j^I < \infty\}} \varepsilon_{(T_j^I, \zeta_j^I)}(ds, dy), \\ \mu^B(ds, dy) &= \sum_{j \geq 1} 1_{\{T_j^B < \infty\}} \varepsilon_{(T_j^B, \zeta_j^B)}(ds, dy). \end{aligned} \tag{6}$$

By our assumptions, under  $\mathcal{Q} := \mathcal{Q}_x^{b, \sigma, \kappa, p, \pi}$ ,  $\mu^I$  is a Poisson random measure on  $(0, \infty) \times \mathbb{R}$  with intensity  $\nu^I(ds, dy) = ds \pi(dy)$ , and  $\mu^B$  is compensated by  $\nu^B(ds, dy)$  given by

$$\nu^B((0, t] \times B) = \int_0^t \varphi_s(1_B \kappa) ds = \int_B \kappa(y) \eta_t(dy), \quad t \geq 0, B \in \mathcal{B} \tag{7}$$

where  $\eta = (\eta_t)_{t \geq 0}$  is the process of occupation time measures under  $\varphi$ :

$$\eta_t(B) = \int_0^t \varphi_s(B) ds, \quad B \in \mathcal{B}, t \geq 0. \tag{8}$$

In the same way we consider point processes of branching times/positions where exactly  $k$  offspring is produced,  $k \neq 1$ :

$$\mu^{B,k}(ds, dy) = \sum_{j \geq 1} 1_{\{T_j^B < \infty, l(\varphi_{(T_j^B)^-}) = l(\varphi_{(T_j^B)^-}) + k - 1\}} \varepsilon_{(T_j^B, \zeta_j^B)}(ds, dy);$$

in particular,  $\mu^{B,0}$  is the point process of death times/positions. We write frequently  $\mu_t^B, \mu_t^I, \dots$  for the measure  $\mu^B((0, t] \times \cdot), \mu^I((0, t] \times \cdot), \dots$  etc.,  $t \geq 0$ .

We define a kernel estimator for the unknown branching rate based on observation of the particle process over a long time interval. For some kernel  $Q$  meeting

$$Q \in \mathcal{C}_{\mathcal{X}}(\mathbb{R}), \int_{\mathbb{R}} Q(y)dy = 1$$

we define with notation of (6) the kernel estimator for the unknown  $\kappa$  at point  $a$  with bandwidth proportional to  $h$ , based on observation of  $\varphi$  up to time  $t$ , by

$$K_{t,h}(a) = \frac{\int_{\mathbb{R}} \frac{1}{h} Q\left(\frac{y-a}{h}\right) \mu_t^B(dy)}{\int_{\mathbb{R}} \frac{1}{h} Q\left(\frac{y-a}{h}\right) \eta_t(dy)}, \quad \text{where } \frac{0}{0} := 0. \tag{9}$$

For small  $h$ ,  $K_{t,h}(a)$  is a natural estimate for  $\kappa(a)$ , for two reasons: first, by (7) and (8),  $(\mu_t^B - \kappa \eta_t)_{t \geq 0}$  (with  $\kappa \eta_t$  the random measure  $\kappa(y) \eta_t(dy)$  on  $\mathbb{R}$ ) is a martingale measure; second, by the ratio limit theorem and definition of  $\bar{m}$ , we have  $P$ -a.s.  $(1/t) \eta_t(f) \rightarrow \bar{m}(f)$  and  $(1/t) \mu_t^B(f) \rightarrow (\kappa \bar{m})(f)$  as  $t \rightarrow \infty$  for fixed  $f$  non-negative, measurable,  $\bar{m}$ -integrable.

2.3. Statistical assumptions

Let  $l(x)$  be the length of the configuration  $x \in S$ . We assume

**Assumption 4** ( $\kappa, p$ )

For all  $q \geq 1$ , we have  $\int_S l^q(x) m(dx) < \infty$ .

**Assumption 5** ( $\kappa$ )

$\int_{\mathbb{R}} y^2 \pi(dy) < \infty$ , and  $\int_{\mathbb{R}} |y|^{2+\delta} \bar{m}(dy) < \infty$  for some  $\delta > 0$ .

In order to introduce our last assumptions, we need some more notation: write  $\varphi^{(t,x)}$  for a subprocess of  $\varphi$  which starts at time  $t$  with one particle located in  $x$ , and which consists of the trajectories followed after time  $t$  by this ancestor and by all its direct descendants up to the time  $D^{(t,x)}$  where the last member of the family dies. Necessarily  $D^{(t,x)}$  is a.s. finite since the void configuration  $\Delta$  is a recurrent atom for  $\varphi$ . Let  $\tau_1^{(t,x)}$  and  $\zeta_1^{(t,x)}$  denote the time and the location of the first branching event in the subprocess  $\varphi^{(t,x)}$ . We introduce first moment measure  $V(\cdot, \cdot)$ , a kernel on  $(\mathbb{R}, \mathcal{B})$ , and second moment measure  $V_2(\cdot, \cdot)$ , a kernel from  $(\mathbb{R}, \mathcal{B})$  to  $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$  (e.g. see Daley & Vere-Jones, 1988, p. 190), for the occupation time measure induced by the subprocess  $\varphi^{(t,x)}$ , given by

$$V(x, B_1) := E \left( \int_t^{D^{(t,x)}} \varphi_s^{(t,x)}(B_1) ds \right),$$

$$V_2(x, B_1 \times B_2) := E \left( \int_t^{D^{(t,x)}} \varphi_s^{(t,x)}(B_1) ds \int_t^{D^{(t,x)}} \varphi_s^{(t,x)}(B_2) ds \right).$$

**Assumption 6** ( $\kappa$ )

For  $f(x) := 1 + |x|$  on  $\mathbb{R}$ , second moments

$$V_2(x, f \otimes f) = E \left( \left( \int_t^{D^{(t,x)}} \varphi_s^{(t,x)}(f) ds \right)^2 \right)$$

are finite for all  $x \in \mathbb{R}$ , and  $V_2(\cdot, f \otimes f)$  belongs to  $L^1(\pi)$  and to  $L^{1+\delta}(\bar{m})$  for some  $\delta > 0$ .

**Assumption 7** ( $k'$ )

We suppose that  $\sigma^2 \in C^{k'+1}$ , bounded away from zero by some constant  $l > 0$ , that  $b \in C^{k'}$ ,  $r, (1 - \varrho) \in C^{k'-1}$ .

2.4. Main results

We describe the quality of  $K_{t,h}$  in the minimax theory. During this section we assume that  $b, \sigma, p, \pi$  are fixed, and that  $\kappa$  is the only parameter of interest. We shall always write  $Q_x^\kappa = Q_x^{b,\sigma,\kappa,p,\pi}$ ,  $E_x^\kappa = E_x^{b,\sigma,\kappa,p,\pi}$ ,  $m^\kappa = m^{b,\sigma,\kappa,p,\pi}$ ,  $Q_m^\kappa = \int_S Q_x^\kappa m^\kappa(dx)$  and  $E_m^\kappa = \int_S E_x^\kappa m^\kappa(dx)$  in order to emphasize the dependence on the unknown  $\kappa$ . With notation  $\beta = k + \alpha$ ,  $k \in \mathbb{N}_0$ ,  $0 < \alpha \leq 1$ , for arbitrary constants  $l > 0$ ,  $r < \infty$ ,  $L$ , consider the following Hölder class

$$H(\beta, l, r, L) = \{ \kappa \in \mathcal{C}^k : l \leq \kappa(x) \leq r, |\kappa^{(k)}(y) - \kappa^{(k)}(x)| \leq L|y - x|^\alpha, \forall x, y \}$$

such a class shapes possible forms of branching rates  $\kappa$ . Fix some  $\nu > 0$  and some compact interval  $I$ . For any norming sequence  $(r_t)_t$  increasing to  $\infty$  and for any process of  $\mathcal{F}_t$ -measurable estimates  $K_t(\cdot)$  for  $\kappa(\cdot)$ , define a squared risk as follows

$$R_t(K_t, H(\beta, l, r, L), r_t) = \sup_{\kappa \in H(\beta, l, r, L)} E_m^\kappa \left\{ r_t^2 \int_I |K_t(a) - \kappa(a)|^2 da | \mathcal{A}_t(\nu, I) \right\} \tag{10}$$

where  $\mathcal{A}_t(\nu, I)$  is roughly an event of type  $\{ \inf_{a \in I} t^{-1} \bar{D}(t, a) \geq \nu \}$  with  $\bar{D}(t, \cdot)$  being the Lebesgue density of  $\eta_t$  (see theorem 5 below). There are several ways to assess the quality of an estimating procedure. Following Hoffmann (1999) and Jacod (2000), we introduce the penalization  $\mathcal{A}_t(\nu, I)$  which imposes that enough information is available on the interval  $I$ . The threshold  $\nu$  gives the minimum amount of information we require. We are able to choose  $\nu$  small enough such that

$$\liminf_{t \rightarrow \infty} \inf_{\kappa \in H(\beta, l, r, L)} Q_m^\kappa(\mathcal{A}_t(\nu, I)) = 1, \tag{11}$$

which shows that intersection with the event  $\mathcal{A}_t(\nu, I)$  is asymptotically no restriction. A precise formulation will be given in lemma 8 and in (59) below. We are now able to state our main results. We denote by  $[x]$  the integer part of  $x \in \mathbb{R}$ .

**Theorem 1**

Grant assumptions 4( $l, \bar{p}$ ), 5( $l$ ), 6( $l$ ) and 7( $[\beta]$ ) with  $\beta \geq 1$ . Assume  $\int Q(y)y^j dy = 0$ ,  $1 \leq j \leq k$  and  $\int Q(y)|y|^\beta dy < \infty$ .

(a) We have for the kernel estimate (9)

$$\limsup_{t \rightarrow \infty} R_t(K_{t,t^{-1/(2\beta+1)}, H(\beta, l, r, L), t^{\beta/(2\beta+1)})} < \infty. \tag{12}$$

(b) For  $h = o(t^{-1/(1+2\beta)})$ , we have for every  $\kappa \in H(\beta, l, r, L)$  and at every point  $a \in \mathbb{R}$

$$\sqrt{th}(K_{t,h}(a) - \kappa(a)) \rightarrow \mathcal{N}(0, \Sigma(a)) \text{ weakly in } \mathbb{R} \text{ under } Q_m^\kappa, \tag{13}$$

where  $\Sigma(a) = \kappa(a)\bar{D}(a)^{-1} \int Q^2(y)dy$ , with  $\bar{D}(\cdot)$  the Lebesgue density of  $\bar{m}$  (see theorem 5).

Our next result shows that the kernel estimate  $K_{t,t^{-1/(2\beta+1)}}$  with largest possible bandwidth  $t^{-1/(2\beta+1)}$  is indeed optimal.

**Theorem 2**

Grant assumptions 4( $l, \bar{p}$ ), 5( $l$ ), 6( $l$ ) and 7(1). Suppose  $\beta \geq 1$ . We have

$$\liminf_{t \rightarrow \infty} \inf_{K_t} R_t(K_t, H(\beta, l, r, L), t^{\beta/(2\beta+1)}) > 0 \tag{14}$$

where the infimum is taken over the class of all estimators.

*Remark 1.* (a) We obtain the same rate  $O(t^{-\beta/(1+2\beta)})$  in theorem 1(a) and theorem 2 as in density estimation in the classical i.i.d. setting.

(b) The condition  $h = o(t^{-1/(1+2\beta)})$  in theorem 2(b) can be generalized to  $h = O(t^{-1/(1+2\beta)})$  by introducing a sequence of bias terms, tight as  $t \rightarrow \infty$ , into the convergence (13). The speed of convergence in (13) is then  $O(t^{-\beta/(1+2\beta)})$  as well.

*Example 1* (Subcritical branching Brownian motion). In subcritical branching Brownian motion, particles move according to independent Brownian motions (i.e.  $b = 0$ ,  $\sigma \equiv 1$  in assumption 1), the branching rate is spatially constant, and the reproduction law—also spatially constant—is a fixed subcritical binary reproduction law ( $\kappa = k$ ,  $k \in (0, \infty)$ ,  $p = \bar{p}$ ,  $\rho < 1$  in assumption 2, with  $p_j > 0$  only for  $j = 0, 2$ ). Suppose moreover that the immigration density  $r$  of assumption 3 is in  $\mathcal{C}_{\mathcal{X}}$ , the class of continuous functions with compact support.

The process  $(l(\varphi_t))_t$  is then a classical subcritical branching process with immigration, and due to the binary structure of reproduction it is easy to see that  $m(\mathbb{R}^l)$  decreases exponentially fast in  $l$  (see Höpfner & Löcherbach, 1999, th. 1.4, Löcherbach, 1999a, sect. 6); so assumption 4 holds.

We can calculate explicitly the Laplace transform of  $\bar{m}$  which exists on a full neighbourhood of 0 (see Höpfner & Löcherbach, 1999, ex. 2.6), so  $\bar{m}$  has moments of arbitrary order, so assumption 5 holds.

For  $f(x) = 1 + |x|$ , we see from lemma 1(c) and the properties of Brownian motion that  $V(0, f^2) < \infty$  and  $V(y, h) = V(0, h(\cdot - y))$ . As a consequence, we have  $V(y, f) = O(|y|)$  and  $V(y, f^2) = O(y^2)$  as  $|y| \rightarrow \infty$ : due to (18) and (19), we obtain  $V_2(y, f \otimes f) = O(y^2)$ , and assumption 6 holds.

So the only remaining assumption for subcritical branching Brownian motion is assumption 7( $k'$ ) requiring that the immigration density  $r$  has  $k'$  continuous derivatives.

*Remark 2.* Thanks to the comparison results which we develop in section 3, the statistical assumptions 4–7 in theorems 1 and 2 are concerning only BD processes with spatially constant branching rate and spatially constant subcritical reproduction law.

*Example 2* (Subcritical BD processes having ergodic particle motion). Consider a BD process with spatially constant branching rate  $\kappa = k$ ,  $k \in (0, \infty)$ , with spatially constant  $p = \bar{p}$  having reproduction mean  $\rho < 1$ , and with ergodic particle motion  $\xi$  in (1). The immigration density  $r$  of assumption 3 is assumed to belong to  $\mathcal{C}_{\mathcal{X}}$ .

As in example 1 before, assumption 4 is a condition on the invariant measure of the continuous time branching process with reproduction law  $p$  and branching rate  $k$ : this invariant measure is required to have moments of arbitrary order. One can check this assumption using classical branching process theory (see Zubkov, 1972; Pakes, 1975, Sewastjanow, 1975; kap. VII.3).

We consider assumption 5 which is a condition concerning essentially the particle motion (1), assumed ergodic. Denote the invariant measure of  $\xi$  in (1) by  $n$ :

$$n(dx) = \frac{2}{\sigma^2(x)} \exp\left(\int_0^x \frac{2b}{\sigma^2}(u)du\right)$$

see e.g. Khasminskii (1980, ex. 2 in sect. III.8). Together with  $\xi$ , also  $\tilde{\xi}$  of lemma 1(b) is ergodic, with invariant measure  $\tilde{n} = kn$ . Necessarily, the immigration measure  $\pi$  has  $\mathcal{C}_{\mathcal{X}}$  density also with respect to  $\tilde{n}$ . Denoting the sup norm of the density  $d\pi/d\tilde{n}$  by  $C$ , (16) and lemma 1(c) show  $\bar{m}(dx) \leq [C/(1 - \varrho)] \tilde{n}(dx) = [Ck/(1 - \varrho)]n(dx)$ . So all integrability conditions with respect to  $\bar{m}$  can be checked from  $n$  which is known explicitly.



In particular, for ergodic Ornstein–Uhlenbeck motion  $\xi$  in (1),  $\bar{m}$  has moments of arbitrary order, and assumption 5 holds.

We consider assumption 6. For  $f(x) = 1 + |x|$ , it is sufficient to prove that  $V(0, f^2) < \infty$ , and  $V(y, f) \leq O(|y|)$  and  $V(y, f^2) \leq O(y^2)$  as  $|y| \rightarrow \infty$ : due to (18) and (19), one deduces assumption 6 from these three properties exactly as in example 1.

In particular, for ergodic Ornstein–Uhlenbeck motion  $\xi$  in (1), we see from the explicit representation of  $\xi$  that assumption 6 holds.

**3. Probabilistic tools**

*3.1. The invariant measure  $\bar{m}$  and comparison results*

**Lemma 1**

(a) *The invariant measure  $\bar{m}$  is given by*

$$\bar{m}(dy) = \int_{\mathbb{R}} \pi(dx) V(x, dy). \tag{15}$$

(b) *Explicit expressions for  $V(\cdot, \cdot)$  can be calculated from*

$$V = \left( \sum_{n=0}^{\infty} (U\kappa\varrho)^n \right) U \quad \pi - a.s., \tag{16}$$

where  $U$  is the occupation time kernel for the diffusion  $d\xi_t = b(\xi_t)dt + \sigma(\xi_t)dW_t$  killed at rate  $\kappa$ , defined by  $U(x, A) = E_x[\int_0^{\infty} 1_A(\xi_t) \exp(-\int_0^t \kappa(\xi_r)dr)dt]$ , and where  $U\kappa\varrho$  is the kernel  $(U\kappa\varrho)(x, dy) := U(x, dy)(\kappa\varrho)(y)$ . This yields the following representation for  $V$ :

$$(V\kappa)(x, B) = E_x \left( \sum_{n=0}^{\infty} \varrho(\tilde{\xi}_{S_1}) \cdots \varrho(\tilde{\xi}_{S_n}) 1_B(\tilde{\xi}_{S_{n+1}}) \right), \tag{17}$$

where  $(V\kappa)$  denotes the kernel  $V(x, dy)\kappa(y)$  and where  $\tilde{\xi}$  is the diffusion  $d\tilde{\xi}_v = (b/\kappa)(\tilde{\xi}_v)dv + (\sigma/\sqrt{\kappa})(\tilde{\xi}_v)d\tilde{W}_v$  with Brownian motion  $\tilde{W}$ . Here,  $S_1, S_2, \dots$  are the jump times of a standard Poisson process independent of  $\tilde{\xi}$ .

(c) *We also have  $(V\kappa)(x, B) = E_x[\int_0^{\infty} 1_B(\tilde{\xi}_t) \exp(-\int_0^t (1 - \varrho)(\tilde{\xi}_s)ds)dt]$ .*

*Proof.* (16) follows by conditioning on the first branching event in the subfamily under consideration, cf. also Etheridge (1993). See Höpfner & Löcherbach (1999) for details concerning (a) and (b) and Höpfner & Löcherbach (2000, rem. 5.28) for (c).

**Lemma 2**

Let  $(V\kappa)_2$  denote the kernel  $(V\kappa)_2(x, (dy, dy')) := V_2(x, (dy, dy'))\kappa(y)\kappa(y')$ . Then

$$(V\kappa)_2(x, B_1 \times B_2) = \int (V\kappa)(x, dz)(I\kappa)_2(z, B_1 \times B_2), \tag{18}$$

where for non-negative  $g, h$ :

$$(I\kappa)_2(z, g \otimes h) = \tau(z)(V\kappa)(z, h)(V\kappa)(z, g) + h(z)(V\kappa)(z, g) + g(z)(V\kappa)(z, h). \tag{19}$$

*Proof.* (18) and (19) follow again by conditioning on the first branching event in a subfamily, using arguments similar to the proof of lemma 1 above.

*Remark 3.* (a) It appears from (17) that the first moment measure  $V(\cdot, \cdot)$  depends on the family  $p(\cdot)$  of reproduction laws only through the first moment function  $\varrho(\cdot)$ .

(b) The representation of lemma 1(c) has the following interpretation:  $(V\kappa)$  is the occupation time kernel for the diffusion  $\xi$  of (17) subject to position dependent killing at rate  $1 - \varrho$ .

(c) From (19) and (a), second moment measures  $V_2(\cdot, \cdot)$  depend on the family of reproduction laws  $p(\cdot)$  only through their first and second moment functions  $\varrho(\cdot)$  and  $\tau(\cdot)$ .

We now develop two comparison results which will be useful in section 4. The main results are theorems 3 and 4, the proofs are obtained via a series of auxiliary results.

**Theorem 3**

For  $\kappa_1 \leq \kappa_2$  and fixed  $b, \sigma, p, \pi$  meeting assumptions 1–3, we have for the first moment kernels

$$V^{(b, \sigma, \kappa_2, p, \pi)}(x, B_1) \leq V^{(b, \sigma, \kappa_1, p, \pi)}(x, B_1) \tag{20}$$

for all  $x \in \mathbb{R}$  and  $B \in \mathcal{B}$ , and thus

$$\bar{m}^{(b, \sigma, \kappa_2, p, \pi)} \leq \bar{m}^{(b, \sigma, \kappa_1, p, \pi)}. \tag{21}$$

If in addition  $\eta := \inf\{\kappa_1/\kappa_2(x) : x \in \mathbb{R}\} > 0$ , we have also

$$V_2^{(b, \sigma, \kappa_2, p, \pi)}(x, B_1 \times B_2) \leq \frac{1}{\eta} V_2^{(b, \sigma, \kappa_1, p, \pi)}(x, B_1 \times B_2). \tag{22}$$

*Proof.* Let  $\varphi$  denote the BD process with parameters  $b, \sigma, \kappa_1, p, \pi$  and  $\varphi^{(t,x)}$  a subprocess starting with one ancestor in  $x$  at time  $t$ . We submit all particles in  $\varphi^{(t,x)}$  to additional position dependent killing at rate  $(\kappa_2 - \kappa_1)(1 - \varrho)$ : together with a particle killed in  $x'$  at time  $t'$  we remove its descendance  $\varphi^{(t',x')}$  from  $\varphi^{(t,x)}$ . Moreover, we introduce fictitious branching points at position dependent rate  $(\kappa_2 - \kappa_1)\varrho$ , where a particle is replaced by exactly one descendant (which we may identify with the parent particle at the fictitious branching time). The resulting thinned processes  $\tilde{\varphi}^{(t,x)}$  resp.  $\tilde{\varphi}$  are BD processes with branching rate  $\kappa_2$  and with position dependent reproduction laws  $\tilde{p}(\cdot)$  given by

$$\tilde{p}_0(x) = \left( \frac{\kappa_1}{\kappa_2} p_0 + \frac{\kappa_2 - \kappa_1}{\kappa_2} (1 - \varrho) \right)(x), \quad \tilde{p}_1(x) = \left( \frac{\kappa_2 - \kappa_1}{\kappa_2} \varrho \right)(x), \quad \tilde{p}_k(x) = \left( \frac{\kappa_1}{\kappa_2} p_k \right)(x), \quad k \geq 2.$$

Note that the first moment function  $\tilde{\varrho}$  of  $\tilde{p}$  coincides with the original  $\varrho$  of  $p$ :

$$\tilde{\varrho}(x) = \sum_{k=1}^{\infty} k \tilde{p}_k(x) = \varrho(x), \tag{23}$$

and that the second moment function  $\tilde{\tau}$  of  $\tilde{p}$  is  $\tilde{\tau}(x) = \sum_{k=2}^{\infty} k(k-1)\tilde{p}_k(x) = (\kappa_1/\kappa_2)(x)\tau(x)$ . For non-negative  $h$  we obtain  $\int_t^{D^{(t,x)}} \tilde{\varphi}_s^{(t,x)}(h) ds \leq \int_t^{D^{(t,x)}} \varphi_s^{(t,x)}(h) ds$  and thus

$$V^{(b, \sigma, \kappa_2, \tilde{p}, \pi)}(x, B_1) \leq V^{(b, \sigma, \kappa_1, p, \pi)}(x, B_1), \tag{24}$$

$$V_2^{(b, \sigma, \kappa_2, \tilde{p}, \pi)}(x, B_1 \times B_2) \leq V_2^{(b, \sigma, \kappa_1, p, \pi)}(x, B_1 \times B_2).$$

Using remark 3(a) and (23), we get (20). Using (20), the structure of  $\tilde{\tau}$  and remark 3(c), we get from (18) and (19)

$$V_2^{b, \sigma, \kappa_2, \tilde{p}, \pi}(x, B_1 \times B_2) \geq \left[ \inf_{x \in \mathbb{R}} \frac{\kappa_1}{\kappa_2}(x) \right] \cdot V_2^{b, \sigma, \kappa_2, p, \pi}(x, B_1 \times B_2)$$

provided

$$\eta := \inf_{x \in \mathbb{R}} \frac{\kappa_1}{\kappa_2}(x) > 0.$$

As a consequence, we get from (24)

$$V_2^{b,\sigma,\kappa_2,p,\pi}(x, B_1 \times B_2) \leq \frac{1}{\eta} V_2^{b,\sigma,\kappa_1,p,\pi}(x, B_1 \times B_2).$$

**Lemma 3**

Write  $\varphi$  for a BD process with parameters  $b, \sigma, \kappa, p, \pi$  meeting assumptions 1–3 and  $\bar{\varphi}$  for a BD process with parameters  $b, \sigma, \kappa, \bar{p}, \pi$  where the reproduction law is the upper bound  $\bar{p}$  of assumption 2. Write  $\bar{R}$  for the life cycle length in  $\bar{\varphi}$  and  $\bar{D}^{(t,x)}$  for the death time of a subprocess  $\bar{\varphi}^{(t,x)}$  of  $\bar{\varphi}$ .

(a) We have for every non-decreasing function  $G : \mathbb{N}_0 \rightarrow \mathbb{R}^+$

$$E_\Delta \left( \int_0^R G(l(\varphi_s)) ds \right) \leq E_\Delta \left( \int_0^{\bar{R}} G(l(\bar{\varphi}_s)) ds \right) \quad \text{and} \quad \frac{1}{c} \leq E_\Delta(R) \leq E_\Delta(\bar{R}). \tag{25}$$

(b) For  $x \in \mathbb{R}, B_1, B_2 \in \mathcal{B}$ , one has

$$\begin{aligned} V^{b,\sigma,\kappa,p,\pi}(x, B_1) &\leq V^{b,\sigma,\kappa,\bar{p},\pi}(x, B_1), \\ V_2^{b,\sigma,\kappa,p,\pi}(x, B_1 \times B_2) &\leq \bar{V}_2^{b,\sigma,\kappa,\bar{p},\pi}(x, B_1 \times B_2). \end{aligned} \tag{26}$$

*Proof.* We embed  $\varphi$  with parameters  $b, \sigma, \kappa, p, \pi$  into a larger BD process  $\bar{\varphi}$ , with parameters  $b, \sigma, \kappa, \bar{p}, \pi$  where the reproduction law is now independent of the space variable in the following way: we generate at each branching event  $(\tau_j^B, \zeta_j^B)$  of  $\varphi$  additional particles according to the reproduction law  $\hat{p}(\zeta_j^B)$  which then migrate and branch as the original particles except that they (and their descendants) use the reproduction law  $\bar{p}$ . Obviously, for every ancestor in position  $x$  at time  $t$ , the original  $\varphi^{(t,x)}$  is embedded as a subprocess into the corresponding  $\bar{\varphi}^{(t,x)}$  formed by the descendance of this ancestor. Thus we have

$$\int_t^{D^{(t,x)}} \varphi_s^{(t,x)}(h) ds \leq \int_t^{\bar{D}^{(t,x)}} \bar{\varphi}_s^{(t,x)}(h) ds, \quad \forall x \in S, \quad \forall t \geq 0 \tag{27}$$

for arbitrary  $h$  non-negative and measurable. (26) now follows immediately from (27). Since we introduce additional particles, we have  $l(\varphi_s) \leq l(\bar{\varphi}_s)$ , thus the first assertion and the main part of the second assertion of (25). The lower bound  $E_\Delta(R) \geq 1/c$  in (25) (with  $c$  the total mass of  $\pi$ ) is trivial since  $E_\Delta(R) > E_\Delta(T_1^l) = 1/c$ .

**Lemma 4**

Write  $\varphi$  for a BD process with parameters  $b, \sigma, \kappa, p, \pi$  meeting assumptions 1–3 and  $\bar{\varphi}$  for a BD process with parameters  $\bar{b}, \bar{\sigma}, l, p, \pi$  where the branching rate is the constant lower bound  $\kappa(\cdot) \geq l > 0$  existing by assumption 2, and where

$$\bar{b} = \frac{l}{\kappa} b, \quad \bar{\sigma} = \sqrt{\frac{l}{\kappa}} \sigma. \tag{28}$$

Write  $\bar{R}$  for the life cycle length in  $\bar{\varphi}$ . Then we have for every non-decreasing function  $G : \mathbb{N}_0 \rightarrow \mathbb{R}^+$

$$E_{\Delta} \left( \int_0^R G(l(\varphi_s)) ds \right) \leq E_{\Delta} \left( \int_0^{\bar{R}} G(l(\bar{\varphi}_s)) ds \right) \quad \text{and} \quad E_{\Delta}(R) \leq E_{\Delta}(\bar{R}). \tag{29}$$

*Proof.* It is possible to “slow down” the BD process  $\varphi$  to get the BD process  $\bar{\varphi}$ : the idea of the transformation is the following. The time change

$$\tau(t) := \inf\{s > 0 : A_s > t\}, \quad A_s := \int_0^s \frac{1}{l} \kappa(\xi_v) dv \tag{30}$$

transforms (see also Höpfner & Löcherbach, 1999c, proof of lem. 1.1) a diffusion  $\xi$  of (1) with position dependent killing at rate  $\kappa(\cdot)$  and with death time denoted by  $\zeta$  into a diffusion  $\bar{\xi}$

$$d\bar{\xi}_t = \bar{b}(\bar{\xi}_t) dt + \bar{\sigma}(\bar{\xi}_t) d\bar{W}_t \tag{31}$$

with killing at constant rate  $l$ , with death time  $\bar{\zeta} = A_{\zeta}$ , and with coefficients given by (28) above. This provides a coupled construction of  $\xi$  killed at rate  $\kappa$  and  $\bar{\xi}$  killed at constant rate  $l$ .

Write  $L(\cdot, \cdot)$  for the local time of  $\xi$ ,  $\bar{L}(\cdot, \cdot)$  for the local time of  $\bar{\xi}$ . By the occupation time formula (e.g. Karatzas & Shreve, 1991, p. 218), a.s.

$$\begin{aligned} \int_{\mathbb{R}} h(a) \frac{2\bar{L}(A_t, a)}{\bar{\sigma}^2(a)} da &= \int_0^{A_t} h(\bar{\xi}_s) ds \geq \int_0^{A_t} h(\bar{\xi}_s) \frac{l}{\kappa(\bar{\xi}_s)} ds \\ &= \int_0^{A_t} h(\xi_{\tau_s}) d\tau_s = \int_0^t h(\xi_v) dv = \int_{\mathbb{R}} h(a) \frac{2L(t, a)}{\sigma^2(a)} da \end{aligned}$$

for every  $h$  non-negative and measurable. Thus the occupation time density  $2\bar{L}(A_t, a)/\bar{\sigma}^2(a)$  of  $\bar{\xi}$  at  $a$  is higher than the occupation time density  $2L(t, a)/\sigma^2(a)$  of  $\xi$  at  $a$ , for every  $a \in \mathbb{R}$  and every  $0 \leq t \leq \zeta$ . Note that local time itself is invariant under this transformation:  $\bar{L}(A_t, a) = L(t, a)$ . Thus a.s.

$$\bar{\zeta} \geq \zeta, \quad \frac{2\bar{L}(\bar{\zeta}, a)}{\bar{\sigma}^2(a)} \geq \frac{2L(\zeta, a)}{\sigma^2(a)} \quad \forall a \in \mathbb{R}. \tag{32}$$

Now we apply this time transformation separately to every particle in the BD process  $\varphi$  between successive branching times. Thus  $\varphi$  is transformed into a BD process  $\bar{\varphi}$  having  $\bar{b}$ ,  $\bar{\sigma}$  of (28) and branching rate  $l$  instead of the original  $b$ ,  $\sigma$ ,  $\kappa$ , whereas  $p$  and  $\pi$  remain unaffected. As a consequence of (32) for every particle,  $l$ -particle configurations have longer life in  $\bar{\varphi}$  than in  $\varphi$ , so (29) holds, and this concludes the proof.

*Remark 4.* Both operations “introducing additional particles” and “slowing down the process” can be combined, and (25) remains true under this combination.

Lemma 3 and 4 yield a second type of comparison result:

**Theorem 4**

Fix some  $b, \sigma, \pi$  and consider a class  $\mathcal{H}$  of parameters  $(\kappa, p)$  such that the BD processes  $\varphi$  corresponding to all pairs  $(b, \sigma, \kappa, p, \pi)$  meet assumptions 1–3, and such that uniformly in  $\mathcal{H}$  lower and upper bounds  $0 < l < r < \infty$  for  $\kappa(\cdot)$  and upper bounds  $\bar{p}$  for  $p(\cdot)$  (in the sense of convolution of probability measures) are available.

(a)

$$\sup_{(\kappa, p) \in \mathcal{H}} E_{\Delta}^{(b, \sigma, \kappa, p, \pi)}(R) < \infty, \quad \inf_{(\kappa, p) \in \mathcal{H}} E_{\Delta}^{(b, \sigma, \kappa, p, \pi)}(R) > 0, \tag{33}$$

and for arbitrary  $G : \mathbb{N}_0 \rightarrow \mathbb{R}^+$  which is non-decreasing,

$$\sup_{(\kappa,p) \in \mathcal{H}} \left( \int_S G(l(x)) m^{(b,\sigma,\kappa,p,\pi)}(dx) \right) \leq cst \left( \int_S G(l(x)) m^{(b,\sigma,l,\bar{p},\pi)}(dx) \right), \tag{34}$$

where the right hand side of (34) does no longer depend on  $b$  and  $\sigma$ .

(b) For  $x \in \mathbb{R}$ ,  $B_1, B_2 \in \mathcal{B}$ , we have

$$\begin{aligned} V^{b,\sigma,\kappa,p,\pi}(x, B_1) &\leq V^{b,\sigma,l,\bar{p},\pi}(x, B_1), \\ V_2^{b,\sigma,\kappa,p,\pi}(x, B_1 \times B_2) &\leq \frac{r}{l} V_2^{b,\sigma,l,\bar{p},\pi}(x, B_1 \times B_2), \end{aligned} \tag{35}$$

and hence

$$\bar{m}^{b,\sigma,\kappa,p,\pi} \leq \bar{m}^{b,\sigma,l,\bar{p},\pi}. \tag{36}$$

(c) For  $Y$  non-negative  $\mathcal{F}_\infty$ -measurable and for all  $(\kappa, p)$  in  $\mathcal{H}$ , we have

$$E_\Delta^{(b,\sigma,\kappa,p,\pi)}(Y) \leq cst E_m^{(b,\sigma,\kappa,p,\pi)}(Y) \tag{37}$$

for some constant only depending on  $l$  and  $\bar{p}$ .

*Proof.* Apply both transformations of lemmata 3 and 4 successively to the original BD process  $\varphi$ ; then (25) and (29) give (34) together with the first assertion of (33)—note that for the BD process with spatially constant branching rate  $l$  and  $\bar{p}$ , the lifetimes of  $l$ -particle configurations are  $\exp(lm)$ -distributed and thus independent of the motion of particles. The second assertion in (33) is the lower bound  $E_\Delta^{(b,\sigma,\kappa,p,\pi)}(R) > 1/c$  of (25). Assertion (35) is theorem 3 together with lemma 3. Assertion (37) is a consequence of (2)

$$E_m(Y) = \int_S m(dx) E_x(Y) \geq m(\{A\}) E_A(Y) = \frac{E_\Delta(T_1^l)}{E_\Delta(R)} E_\Delta(Y) = \frac{1/c}{E_\Delta(R)} E_\Delta(Y)$$

together with the uniform bounds on  $E_\Delta^{(b,\sigma,\kappa,p,\pi)}(R)$  of (33).

### 3.2. Local time for the BD process

Introducing a local time for the particle process  $\varphi$  and studying its asymptotics (with the help of a Tanaka formula for BD processes), we shall prove below the following theorems 5 and 6. From these we shall deduce the asymptotic properties of the kernel estimates. Write  $Q_x = Q_x^\kappa$ ,  $m = m^\kappa$ ,  $Q_m = \int_S Q_x m(dx)$ , and  $\mathcal{A}$  and  $\mathbb{F}$  for the  $\sigma$ -field  $\mathcal{A}$  and the filtration  $\mathbb{F}$  completed with respect to  $Q_m$ . During this subsection, we consider the stationary case and work on the stochastic basis  $(\Omega, \mathcal{A}, \mathbb{F})$  (which satisfies the ‘‘usual hypotheses’’).

#### Theorem 5

Under assumptions 1–3, 4( $\kappa, p$ ) and 5( $\kappa$ ), the invariant measure  $\bar{m}$  has a Lebesgue density  $\bar{D}(\cdot)$  which is continuous and strictly positive on  $\mathbb{R}$ . On  $(\Omega, \mathcal{A}, \mathbb{F}, Q_m)$  there exists a stochastic process  $(\bar{D}(t, \cdot))_{t \geq 0}$  of Lebesgue densities for the occupation time measure process  $(\eta_t)_{t \geq 0}$  such that

$$\sup_{a \in K} E_m \left\{ [t^{-1} \bar{D}(t, a) - \bar{D}(a)]^2 \right\} \rightarrow 0 \quad \text{as } t \rightarrow \infty$$

for arbitrary compact subsets  $K$  of  $\mathbb{R}$ .  $(\bar{D}(t, a))_{t \geq 0, a \in \mathbb{R}}$  has the path properties of semimartingale local time.

For semimartingale local time, we refer the reader to Karatzas & Shreve (1991, p. 218) or Revuz & Yor (1991, ch. VI).

**Theorem 6**

Under assumptions 1–3, 4( $\kappa, p$ ), 5( $\kappa$ ) and 6( $\kappa$ ) we have on  $(\Omega, \tilde{\mathcal{A}}, \tilde{\mathbb{F}})$

$$\sup_{a \in K} \sup_{t \geq 1} E_m \left\{ t \left[ t^{-1} \bar{D}(t, a) - \bar{D}(a) \right]^2 \right\} < \infty \tag{38}$$

for arbitrary compact subsets  $K$  of  $\mathbb{R}$ .

3.2.1. Proofs of theorems 5 and 6

The proofs of theorems 5 and 6 will be done through several steps. We start with introducing the local time for the BD process.

On a stochastic interval  $[[T_{n-1}, T_n[[$ , every particle  $i$ ,  $1 \leq i \leq l(\varphi_{T_{n-1}})$ , travels on a diffusion path. So there is a semimartingale local time associated to every particle and its random life time. Pasting this together (the technical details are exactly analogous to the definition of local time for “birth and death on a flow” models considered in Höpfner & Löcherbach (1998) and are omitted here), we obtain a local time for the particle process  $\varphi$ :

**Lemma 5**

On  $(\Omega, \tilde{\mathcal{A}}, \tilde{\mathbb{F}})$ , a local time  $(\bar{L}(t, a))_{t \geq 0, a \in \mathbb{R}}$  for  $\varphi$  exists, satisfying

- (i)  $t \rightarrow \bar{L}(t, a)$  is non-decreasing with  $\bar{L}(0, a) \equiv 0$  for all  $a \in \mathbb{R}$ .
- (ii)  $(a, t) \rightarrow \bar{L}(t, a)$  is bi-continuous.
- (iii) For  $f$  non-negative and measurable, we have  $Q_m$ -a.s.

$$\int_0^t \varphi_s(f) ds = \int_{\mathbb{R}} f(a) \frac{2\bar{L}(t, a)}{\sigma^2(a)} da, \quad \forall t \geq 0. \tag{39}$$

As a consequence, for  $Q_m$ -almost all  $\omega \in \Omega$ , the occupation time measures  $\eta_t(\omega)$ ,  $t \geq 0$ , admit Lebesgue densities

$$a \rightarrow \bar{D}(t, a)(\omega) := \frac{2\bar{L}(t, a)(\omega)}{\sigma^2(a)}, \quad t \geq 0.$$

Proceeding in analogy to Höpfner & Löcherbach (1998) for the BD process, we obtain a Tanaka formula.

**Lemma 6**

(a) Let  $M^{a,1}$  denote the continuous local  $\mathbb{F}$ -martingale

$$dM_t^{a,1} = \sum_{i=1}^{l(\varphi_{T_{n-1}})} \text{sgn}(\varphi_{t-}^i - a) \sigma(\varphi_{t-}^i) dW_t^i \quad \text{on } [[T_{n-1}, T_n[[, n \geq 1; M_0^{a,1} \equiv 0.$$

On  $(\Omega, \tilde{\mathcal{A}}, \tilde{\mathbb{F}})$ , for  $a \in \mathbb{R}$  fixed, we have, up to  $Q_m$ -indistinguishability,

$$\begin{aligned} & \varphi_t(|\cdot - a|) - \varphi_0(|\cdot - a|) - |\cdot - a| * \mu_t^I - \sum_k (k-1) (|\cdot - a| * \mu^{B,k})_t \\ &= \int_0^t \varphi_{s-}(\text{sgn}(\cdot - a)b) ds + 2\bar{L}(t, a) + M_t^{a,1}. \end{aligned} \tag{40}$$

(b) The one-sided variant of the Tanaka formula is

$$\begin{aligned} & \varphi_t((\cdot - a)^-) - \varphi_0((\cdot - a)^-) - (\cdot - a)^- * \mu_t^I - \sum_k (k - 1)((\cdot - a)^- * \mu^{B,k})_t \\ &= - \int_0^t \varphi_s(1_{(-\infty, a)}b)ds + \bar{L}(t, a) + \tilde{M}_t^{a,1} \end{aligned} \tag{41}$$

up to  $\mathcal{Q}_m$ -indistinguishability, with  $\tilde{M}^{a,1}$  the continuous local  $\mathbb{F}$ -martingale given by

$$\tilde{M}_0^{a,1} \equiv 0 \quad \text{and} \quad d\tilde{M}_t^{a,1} = - \sum_{i=1}^{I(\varphi_{T_{n-1}})} 1_{(-\infty, a)}(\varphi_{t-}^i)\sigma(\varphi_{t-}^i)dW_t^i \quad \text{on } [[T_{n-1}, T_n[[, \quad n \geq 1.$$

Asymptotics of local time follow from the Tanaka formula. This has been noticed by Kutoyants (1998) for classical one-dimensional ergodic diffusions. First note that  $\langle M^{a,1} \rangle_t = \int_0^t \varphi_s(\sigma^2)ds$ , so the Lipschitz condition on  $\sigma$  together with finiteness of  $\int_{\mathbb{R}} (1 + |y|)^2 \bar{m}(dy)$  by assumption 5( $\kappa$ ) implies

$$\sup_{t>0} \sup_{a \in K} t^{-1} E_m(\langle M^{a,1} \rangle_t) < \infty. \tag{42}$$

Since  $\int_{\mathbb{R}} y^2 \pi(dy)$  is finite by assumption 5( $\kappa$ ), since second moments  $\sum_k (k - 1)^2 p_k(\cdot)$  are bounded on  $\mathbb{R}$  by assumption 2, also

$$M_t^{a,2} = |\cdot - a| * \mu_t^I - t \int_{\mathbb{R}} |\cdot - a| d\pi$$

and

$$M_t^{a,3} = \sum_k (k - 1)(|\cdot - a| * \mu^{B,k})_t - \int_0^t \varphi_s((\varrho - 1)\kappa|\cdot - a|)ds$$

are  $\mathbb{F}$ -martingales meeting

$$\sup_{t>0} \sup_{a \in K} t^{-1} E_m(\langle M^{a,i} \rangle_t) < \infty, \quad i = 2, 3. \tag{43}$$

Next, we define

$$\Lambda(t, a) = \int_0^t \varphi_s((1 - \varrho)\kappa|\cdot - a| - \text{sgn}(\cdot - a)b)ds$$

and

$$\Lambda(a) = \int_{\mathbb{R}} ((1 - \varrho)\kappa|\cdot - a| - \text{sgn}(\cdot - a)b)d\bar{m}.$$

$\kappa$  being bounded by assumption 2 and  $b$  Lipschitz,  $\Lambda$  is a continuous function on  $\mathbb{R}$  (assumption 5( $\kappa$ ) and dominated convergence). Finally, introduce  $\bar{L}$  (continuously by assumption 5( $\kappa$ )) by

$$2\bar{L}(a) = \Lambda(a) - \int_{\mathbb{R}} |\cdot - a| d\pi. \tag{44}$$

By the Tanaka formula (40), the random variable  $2\sqrt{t}[t^{-1}\bar{L}(t, a) - \bar{L}(a)]$  can be rewritten as

$$\frac{(\varphi_t - \varphi_0)(|\cdot - a|)}{\sqrt{t}} - \frac{M_t^{a,1} + M_t^{a,2} + M_t^{a,3}}{\sqrt{t}} + \sqrt{t} \left( \frac{\Lambda(t, a)}{t} - \Lambda(a) \right). \tag{45}$$

The form (45) of the Tanaka formula is the key to the proofs of theorems 5 and 6.

*Proof of theorem 5.* Since  $E_m(\Lambda(t, a)) = t\Lambda(a)$ , we see that  $E_m(\bar{L}(t, a)) = t\bar{L}(a)$ , for all  $a \in \mathbb{R}$ . Thus we can take expectations in (41) to see that

$$a \rightarrow \bar{D}(a) := \frac{2\bar{L}(a)}{\sigma^2(a)} \tag{46}$$

is a continuous density of the invariant measure  $\bar{m}$ . This density is also strictly positive—this follows from the representation of  $\bar{m}$  in lemma 1 together with the fact that 1-potentials of diffusions with strictly positive diffusion coefficient have strictly positive Lebesgue densities. Next, assumption 4( $\kappa, p$ ) and assumption 5( $\kappa$ ) imply

$$E_m\left((\varphi_1(1 + |\cdot|))^2\right) < \infty. \tag{47}$$

This follows from Jensen inequality for  $\eta > 1$

$$\left(\left(\frac{1}{I(\varphi_1)}\varphi_1\right)(1 + |\cdot|)\right)^\eta \leq \left(\frac{1}{I(\varphi_1)}\varphi_1\right)((1 + |\cdot|)^\eta) \leq \varphi_1((1 + |\cdot|)^\eta)$$

on  $\{I(\varphi_1) > 0\}$  and from Hölder inequality with

$$\frac{1}{p} + \frac{1}{q} = 1, \quad q = 1 + \frac{\delta}{2}$$

$$E_m\{(\varphi_1(1 + |\cdot|))^2\} \leq [E_m\{(I(\varphi_1))^{2p}\}]^{1/p} [E_m\{\varphi_1((1 + |\cdot|)^{2+\delta})\}]^{1/q}$$

together with assumption 4( $\kappa, p$ ) and assumption 5( $\kappa$ ). Due to (47), we have for compact  $K$ :

$$\sup_{a \in K} \sup_{t \geq 0} E_m\{(\varphi_t(|\cdot - a|))^2\} < \infty. \tag{48}$$

By exactly the same type of argument, we see that the family  $\{(t^{-1}\Lambda(t, a))^2: t \geq 0, a \in K\}$  is uniformly integrable under  $\mathcal{Q}_m$ . For  $a \in \mathbb{R}$  fixed, the ratio limit theorem shows

$$t^{-1}\Lambda(t, a) \rightarrow \Lambda(a) \quad \mathcal{Q}_m\text{-a.s. as } t \rightarrow \infty.$$

We consider joint convergence  $(a', t) \rightarrow (a, \infty)$ ,  $a \in \mathbb{R}$ . Since  $\bar{m}$  has a continuous Lebesgue density, limits as  $t \rightarrow \infty$  for  $t^{-1} \int_0^t \varphi_s(1_{(a-\varepsilon, a+\varepsilon)} + \varepsilon) ds$  exist  $\mathcal{Q}_m$ -a.s., and can be made arbitrarily small by small choice of  $\varepsilon$ . Using bounds of type  $\int_0^t \varphi_s(1_{(a-\varepsilon, a+\varepsilon)} + \varepsilon) ds$  for differences  $\Lambda(t, a') - \Lambda(t, a)$  we get

$$t^{-1}\Lambda(t, a') \rightarrow \Lambda(a) \quad \mathcal{Q}_m\text{-a.s. as } (a', t) \rightarrow (a, \infty).$$

Combining this with uniform integrability, we have

$$\sup_{a \in K} E_m\{(t^{-1}\Lambda(t, a) - \Lambda(a))^2\} \rightarrow 0 \quad \text{as } t \rightarrow \infty. \tag{49}$$

Dividing both sides of the Tanaka formula (45) by  $\sqrt{t}$ , (42), (43), (48) and (49) show

$$\sup_{a \in K} E_m\{(t^{-1}\bar{L}(t, a) - \bar{L}(a))^2\} \rightarrow 0 \quad \text{as } t \rightarrow \infty. \tag{50}$$

*Proof of theorem 6.* By the Tanaka formula (45) and (42), (43) and (48), it remains to show that (49) can be strengthened under the additional assumption 6( $\kappa$ ) to

$$\sup_{a \in K} \sup_{t \geq 1} E_m\{t(t^{-1}\Lambda(t, a) - \Lambda(a))^2\} < \infty. \tag{51}$$

(1) In a first step, we reformulate the condition assumption 6( $\kappa$ ) in order to prove (51). Let  $\varphi^{(I, j)}$  denote the subprocess formed by all particles stemming from the ancestor who immigrated at time  $T_j^I$ , and let  $D_j^I$  denote the time where the last member of this family dies.



Similarly, let  $\varphi^{(O)}$  denote the subprocess consisting of all particles which are descendants of particles in the original configuration  $\varphi_0^{(O)} = \varphi_0$ , and write  $D^O$  for the death time of the last member of one of these families. We show that assumption 6( $\kappa$ ) implies the following:

**Assumption 6'**

$$E \left\{ \left( \int_{T_j^D}^{D_j^D} \varphi^{(I,j)}(1 + |\cdot|) ds \right)^2 \right\} < \infty, E_m \left\{ \left( \int_0^{D^O} \varphi^{(O)}(1 + |\cdot|) ds \right)^2 \right\} < \infty.$$

In fact, the first assertion of assumption 6' is

$$E \left\{ \left( \int_{T_j^D}^{D_j^D} \varphi^{(I,j)}(1 + |\cdot|) ds \right)^2 \right\} = \int \frac{1}{c} \pi(dy) V_2(y, f \otimes f) < \infty$$

with  $f(x) \equiv 1 + |x|$ , guaranteed by assumption 6( $\kappa$ ). We turn to the second one. Here

$$\begin{aligned} E_m \left\{ \left( \int_0^{D^O} \varphi^{(O)}(1 + |\cdot|) ds \right)^2 \right\} &= E_m \left\{ \sum_{i=1}^{l(\varphi_0)} (V_2(\varphi_0^i, f \otimes f) - V^2(\varphi_0^i, f)) + \sum_{i,j=1}^{l(\varphi_0)} V(\varphi_0^i, f) V(\varphi_0^j, f) \right\} \\ &= \int_{\mathbb{R}} d\bar{m} (V_2(\cdot, f \otimes f) - V^2(\cdot, f)) + \int_S m(dx) \left( \sum_{i=1}^{l(x)} V(x^i, f) \right)^2. \end{aligned}$$

By definition of  $V_2$  as a second moment, we have certainly  $V_2(\cdot, f \otimes f) - V^2(\cdot, f) \geq 0$ . Note that

$$\left[ \sum_{i=1}^{l(x)} V(x^i, f) \right]^2 \leq l^2(x) \frac{1}{l(x)} \sum_{i=1}^{l(x)} V^2(x^i, f) \leq l^2(x) \frac{1}{l(x)} \sum_{i=1}^{l(x)} V_2(x^i, f \otimes f)$$

and use Hölder inequality as in the proof of theorem 5: then assumption 4( $\kappa, p$ ) and assumption 6( $\kappa$ ) guarantee that also the second integral in assumption 6' is finite. So assumption 6' is proved.

(2) We introduce some more notations. For measurable  $f$  meeting  $|f| \leq \text{cst}(1 + |\cdot|)$  write

$$H_t(f) = \int_0^t (\varphi_s(f) - \bar{m}(f)) ds.$$

For the particular choice  $f = f_a := (1 - \varrho)\kappa|\cdot - a| - \text{sgn}(\cdot - a)b$  and  $a \in K$ , the processes

$$H_t(f_a) = A(t, a) - tA(a), \quad t \geq 0$$

are our object of interest. Adjoining to  $H_t(f)$  the “future” of families with origin before time  $t$  we define (with notations of assumption 6') new processes

$$K_t(f) = \sum_{j \geq 1} 1_{\{T_j^t \leq t\}} \int_{T_j^t}^{D_j^D} \varphi_s^{(I,j)}(f) ds - t\bar{m}(f), \quad t \geq 0$$

and

$$G_t(f) = \int_0^{D^O} \varphi_s^{(O)}(f) ds + K_t(f), \quad t \geq 0$$

(both of pure jump type, with internal filtration bigger than  $\tilde{\mathbb{F}}$ ). Clearly  $H_t(f) \leq G_t(f)$  for all  $t \geq 0$ . By the basic independence assumptions in assumptions 1–3,  $(K_t(f))_{t \geq 0}$  is a compound Poisson process with i.i.d. marks  $\int_{T_j^t}^{D_j^D} \varphi_s^{(I,j)}(f) ds, j \geq 1$ , and independent exponential waiting times with parameter  $c$  (note that  $\bar{m} = \pi V$ , see lemma 1, and that the expected jump height in  $(K_t(f))_t$  under  $\mathcal{Q}_m$  is  $\bar{m}(f)/c$ ). By assumption 6', second moments exist for the jumps, thus

$$\sup_{a \in K} \sup_{t \geq 1} E_m \{ t^{-1} K_t^2(f_a) \} < \infty. \tag{52}$$

Next,  $\int_0^{D^0} \varphi_s^{(O)}(f_a) ds = G_t(f_a) - K_t(f_a)$  is independent of  $t$ , so by assumption 6'

$$\sup_{a \in K} \sup_{t \geq 1} E_m \{ (G_t(f_a) - K_t(f_a))^2 \} < \infty. \tag{53}$$

Write  $\vartheta_t$  for the shift operator on  $(\Omega, \mathcal{A})$ . We have for all  $\omega$

$$(G_t(f_a) - H_t(f_a))(\omega) = \left( \left[ \int_0^{D^0} \varphi_s^{(O)}(f_a) ds \right] \circ \vartheta_t \right)(\omega)$$

so Markov property and stationarity of  $\varphi$  show

$$\sup_{a \in K} \sup_{t \geq 1} E_m \{ [G_t(f_a) - H_t(f_a)]^2 \} < \infty. \tag{54}$$

We can conclude, since combining (52)–(54), we have

$$\sup_{a \in K} \sup_{t \geq 1} E_m \left\{ \left[ t^{-1/2} H_t(f_a) \right]^2 \right\} < \infty. \tag{55}$$

**Corollary 1**

Grant assumptions 4(l,  $\bar{p}$ ), 5(l) and 6(l). Then we have

$$\sup_{\kappa \in H(\beta, l, r, L)} \sup_{a \in K} \sup_{t \geq 1} E_m^\kappa \left\{ t [t^{-1} \bar{D}(t, a) - \bar{D}^\kappa(a)]^2 \right\} < \infty \tag{56}$$

for arbitrary compact subsets  $K$  of  $\mathbb{R}$ .

*Proof of corollary 1.* Note that by our assumptions and by theorem 4(b),

$$\sup_{\kappa \in H(\beta, l, r, L)} \int |y|^{2+\delta} \bar{m}^\kappa(dy) = \int |y|^{2+\delta} \bar{m}^l(dy) < \infty;$$

thus (42)–(43) and (48) hold uniformly in  $\kappa$ . As a consequence, it suffices to show that (51) holds uniformly in  $\kappa$  which is shown analogously to the proof of theorem 6 — using the fact that due to theorem 4(b) assumption 6(l) implies that assumption 6 also holds uniformly in  $\kappa \in H(\beta, l, r, L)$ .

3.2.2. *Comments on the invariant density*

As a direct consequence of (46), we know that  $\bar{D}^\kappa$  is given by  $\bar{D}^\kappa(a) = 2\bar{L}(a)/\sigma^2(a)$  with  $\bar{L}(\cdot)$  continuous and non-negative on  $\mathbb{R}$ . Note that  $\bar{L} = \bar{L}^\kappa$  depends on  $\kappa$ . Taking expectations in the one-sided Tanaka formula (41) and using dominated convergence,  $\bar{L}$  satisfies the equation

$$\bar{L}^l(a) = - \int_{-\infty}^a r(x) dx + \int_{-\infty}^a \kappa(x)(1 - \varrho)(x) \frac{2\bar{L}}{\sigma^2}(x) dx + b(a) \frac{2\bar{L}}{\sigma^2}(a). \tag{57}$$

If  $\sigma^2$  is  $C^2$  and  $b$  is  $C^1$ , then  $\bar{D} = 2\bar{L}/\sigma^2$  is  $C^2$ , and (57) can be written as

$$A^* \bar{D}^\kappa - \kappa(1 - \varrho) \bar{D}^\kappa = -r \tag{58}$$

where

$$A^* u := \left( \frac{\sigma^2}{2} u \right)'' - (bu)'$$

is the adjoint of the generator  $A$  of the diffusion of assumption 1.

**Lemma 7**

Suppose that assumption  $7([\beta])$  holds. Then for all  $\kappa \in H(\beta, l, r, L)$  with  $\beta = k + \alpha$  also  $\bar{D}^\kappa$  belongs to some Hölder ball  $H(\beta, l', r', L')$ , with  $l', r', L'$  independent of  $\kappa$ .

*Proof.* Note that derivatives  $\kappa^{(k')}$  for  $1 \leq k' \leq k$  are uniformly bounded for all  $\kappa \in H(\beta, l, r, L)$ . Then the assertion follows inductively in  $k$  from (57).

**4. Proofs**

*4.1. Preliminaries*

This section is devoted to the proof of theorems 1 and 2. The proofs for the lemmata we need are found in the appendix.

For  $a \in \mathbb{R}$ , write  $A^{t,h}(a) = A^{t,h} := t^{-1} \int_{\mathbb{R}} h^{-1} Q(h^{-1}(y - a)) \eta_t(dy)$ . We now give the explicit definition of  $\mathcal{A}_t(v, I)$  appearing in (10):

$$\mathcal{A}_t(v, I) := \bigcap_{a \in I} \{A^{t,h_t}(a) \geq v\}. \tag{59}$$

Here we use the bandwidth  $h_t = t^{-1/(2\beta+1)}$ , a choice which proves to be technically most suitable for us. Generalizing (11) we have:

**Lemma 8**

Under the assumptions of theorem 2,

$$\liminf_{t \geq 1} \inf_{\kappa \in H(\beta, l, r, L)} Q_m^\kappa \left\{ \inf_{a \in I} A^{t,h_t}(a) \geq v \right\} = 1$$

for all  $v < \inf_{\kappa \in H(\beta, l, r, L)} \inf_{a \in I} \bar{D}^\kappa(a)$ , for all sequences  $h_t \rightarrow 0$  with  $th_t \rightarrow \infty$  as  $t \rightarrow \infty$ .

*4.2. Proof of theorem 1*

Throughout this subsection, the stochastic basis is the statistically relevant  $(\Omega, \mathcal{A}, \mathbb{F})$ , with the stationary law  $Q_m$ . We note

$$M^{t,h} := M^{t,h}(a) := \frac{1}{\sqrt{t}} \int_{\mathbb{R}} \frac{1}{\sqrt{h}} Q\left(\frac{y-a}{h}\right) (\mu_t^B - \nu_t^B)(dy). \tag{60}$$

The proof for theorem 1 is based on the decomposition

$$A^{t,h}(K_{t,h}(a) - \kappa(a)) = \frac{M^{t,h}}{\sqrt{th}} + \frac{1}{t} \int \frac{1}{h} Q\left(\frac{y-a}{h}\right) (\kappa(y) - \kappa(a)) \eta_t(dy) \tag{61}$$

of the estimation error of the kernel estimator.

**Lemma 9**

Under assumptions 1–3,  $4(\kappa, p)$  and  $5(\kappa)$ , for continuous functions  $f$  on  $\mathbb{R}$ , we have

$$\frac{1}{t} \int_{\mathbb{R}} \frac{1}{h} Q\left(\frac{y-a}{h}\right) f(y) \eta_t(dy) \rightarrow f(a) \bar{D}(a)$$

in  $L^2(Q_m)$  as  $t \rightarrow \infty$  and  $h \rightarrow 0$ , and thus in  $Q_m$ -probability.

**Lemma 10**

Under assumptions 1–3,  $4(\kappa, p)$  and  $5(\kappa)$ , if  $t \rightarrow \infty$ ,  $h \rightarrow 0$ ,  $th \rightarrow \infty$ ,

$$\frac{1}{\sqrt{t}} \int_{\mathbb{R}} \frac{1}{\sqrt{h}} \mathcal{Q}\left(\frac{y-a}{h}\right) (\mu_t^B - \nu_t^B)(dy) \rightarrow \mathcal{N}(0, \kappa(a) \int \mathcal{Q}^2(y) dy \bar{D}(a))$$

weakly under  $\mathcal{Q}_m$  as  $t \rightarrow \infty$ .

Now we are able to give the proof of theorem 1.

*Proof of theorem 1(b).* Expanding  $\kappa$  we have

$$\kappa(y) - \kappa(a) = \sum_{j=1}^k \frac{\kappa^{(j)}(a)}{j!} (y-a)^j + \frac{\kappa^{(k)}(u) - \kappa^{(k)}(a)}{k!} (y-a)^k \tag{62}$$

where  $u$  is between  $y$  and  $a$ . We consider separately the contributions to (61) coming from terms in this expansion.

(1) The last derivative  $\kappa^{(k)}$  being Hölder  $\alpha$ , lemma 9 shows

$$\frac{1}{t} \int \frac{1}{h} \mathcal{Q}\left(\frac{y-a}{h}\right) \left(\frac{\kappa^{(k)}(u) - \kappa^{(k)}(a)}{k!} (y-a)^k\right) \eta_t(dy) = o_{\mathcal{Q}_m}(h^\beta) \tag{63}$$

as  $t \rightarrow \infty$ ,  $h \rightarrow 0$ . For  $h = o(t^{-1/(1+2\beta)})$  the expression (63) is thus  $o_{\mathcal{Q}_m}(1/\sqrt{th})$  as desired.

(2) Proceeding as in the proof of lemma 9, we combine theorem 6 and Jensen inequality to see that for  $j = 1, 2, \dots, k$ ,

$$E_m \left\{ \left( \int_{\mathbb{R}} \frac{1}{h} \mathcal{Q}\left(\frac{y-a}{h}\right) (y-a)^j \sqrt{t} \left(\frac{\eta_t}{t} - \bar{m}\right)(dy) \right)^2 \right\} = O(h^{2j}) = o(1)$$

as  $t \rightarrow \infty$ ,  $h \rightarrow 0$ . Thus

$$\int_{\mathbb{R}} \frac{1}{h} \mathcal{Q}\left(\frac{y-a}{h}\right) (y-a)^j \left(\frac{\eta_t}{t} - \bar{m}\right)(dy) = o_{\mathcal{Q}_m}\left(\frac{1}{\sqrt{t}}\right) \tag{64}$$

for  $j = 1, \dots, k$ , as  $t \rightarrow \infty$  and  $h \rightarrow 0$ ; *a fortiori*, (64) is  $o_{\mathcal{Q}_m}(1/\sqrt{th})$ .

(3) For  $j = 1, \dots, k$ , only deterministic terms

$$\int \frac{1}{h} \mathcal{Q}\left(\frac{y-a}{h}\right) (y-a)^j \bar{m}(dy) = O(h^{k+1}) \tag{65}$$

remain (the density  $\bar{D}$  of  $\bar{m}$  being  $\mathcal{C}^k$ , (65) combines an expansion of  $\bar{D}$  at  $a$  with the properties of  $\mathcal{Q}$ ). Again for  $h = o(t^{-1/(1+2\beta)})$ , the expression (65) is  $o_{\mathcal{Q}_m}(1/\sqrt{th})$  as desired.

(4) We have seen via (62) and (1)–(3) above that the bias term on the r.h.s of (61) is  $o_{\mathcal{Q}_m}(1/\sqrt{th})$  under our conditions: so theorem 1(b) is proved.

*Remark 5.* Replacing small bandwidth  $h = o(t^{-1/(2\beta+1)})$  by  $h = O(t^{-1/(2\beta+1)})$ , we get

$$\frac{1}{t} \int \frac{1}{h} \mathcal{Q}\left(\frac{y-a}{h}\right) (\kappa(y) - \kappa(a)) \eta_t(dy) = O_{\mathcal{Q}_m}\left(\frac{1}{\sqrt{th}}\right) \tag{66}$$

and thus obtain asymptotic normality of  $\sqrt{th}(K_{t,h}(a) - \kappa(a))$  up to a tight sequence of bias terms.

*Proof of theorem 1(a).* The proof is analogous to the proof of 1(b) above. Using the representation (61) of estimation errors, we get for  $r_t = t \cdot h_t$  with  $h_t = t^{-1/(2\beta+1)}$ ,

$$E_m^k \left\{ r_t^2 \int_I (K_{t,t^{-1/(2\beta+1)}}(a) - \kappa(a))^2 da | \mathcal{A}_t(v, I) \right\} \leq \frac{2}{v^2} [I_1 + I_2]$$

where

$$I_1 = E_m^\kappa \left\{ \int_I M^{t,h}(a)^2 da | \mathcal{A}_t(v, I) \right\}$$

and

$$I_2 = E_m^\kappa \left\{ \int_I \left[ \frac{1}{r_t} \int \mathcal{Q} \left( \frac{y-a}{h_t} \right) (\kappa(y) - \kappa(a)) \eta_t(dy) \right]^2 da | \mathcal{A}_t(v, I) \right\}.$$

(1) We deal with the martingale term  $I_1$  :

$$\begin{aligned} I_1 &\leq \frac{1}{\mathcal{Q}_m^\kappa(\mathcal{A}_t(v, I))} \cdot \int_I E_m^\kappa \left\{ \frac{1}{th_t} \int \mathcal{Q}^2 \left( \frac{y-a}{h_t} \right) \kappa(y) \bar{D}_t(y) dy \right\} da \\ &\leq \frac{M}{\mathcal{Q}_m^\kappa(\mathcal{A}_t(v, I))} \left( \int_I E_m^\kappa \left\{ \int \mathcal{Q}^2(x) \frac{\bar{D}_t}{t} (a + h_t x) dx \right\} da \right). \end{aligned}$$

Note that by theorem 4 and corollary 1,

$$\sup_{\kappa \in H(\beta, l, r, L)} \sup_{t \geq 1} \sup_{a \in K} E_m^\kappa \{ t^{-1} \bar{D}_t(a) \} < \infty,$$

thus together with lemma 8,  $\limsup_{t \geq 1} \sup_{\kappa \in H(\beta, l, r, L)} I_1$  is finite.

(2) We write

$$I_2 \leq \frac{1}{\mathcal{Q}_m^\kappa(\mathcal{A}_t(v, I))} \cdot E_m^\kappa \left\{ \int_I [A_t(a)]^2 da \right\}$$

where

$$\begin{aligned} A_t(a) &= \frac{t}{r_t} \int_{\mathbb{R}} \mathcal{Q} \left( \frac{y-a}{h_t} \right) (\kappa(y) - \kappa(a)) \left( \frac{\bar{D}_t}{t}(y) - \bar{D}^\kappa(y) \right) dy \\ &\quad + \frac{t}{r_t} \int_{\mathbb{R}} \mathcal{Q} \left( \frac{y-a}{h_t} \right) (\kappa(y) - \kappa(a)) \bar{D}^\kappa(y) dy. \end{aligned}$$

In a first step, thanks to the uniform Lipschitz continuity of the functions  $\kappa$  with, say, Lipschitz constant  $L$ , we arrive at

$$\begin{aligned} E_m^\kappa \left\{ \int_I \left[ \frac{t}{r_t} \int_{\mathbb{R}} \mathcal{Q} \left( \frac{y-a}{h_t} \right) (\kappa(y) - \kappa(a)) \left( \frac{\bar{D}_t}{t}(y) - \bar{D}^\kappa(y) \right) dy \right]^2 da \right\} \\ \leq L^2 h_t^3 \sup_{a \in K} \sup_{t \geq 1} \sup_{\kappa \in H(\beta, l, r, L)} E_m^\kappa \left( t \left( \frac{\bar{D}_t}{t} - \bar{D}^\kappa \right)^2 \right) \cdot \int_{\mathbb{R}} \mathcal{Q}(x) dx \end{aligned}$$

for a suitable compact  $K$ , and this expression remains bounded as  $t \rightarrow \infty$  due to corollary 1.

It remains to consider

$$E_m^\kappa \left\{ \int_I \left( \frac{t}{r_t} \int \mathcal{Q} \left( \frac{y-a}{h_t} \right) (\kappa(y) - \kappa(a)) \bar{D}^\kappa(y) dy \right)^2 da \right\}.$$

Here one uses the properties of the kernel  $\mathcal{Q}$  and the fact that the set of densities  $\{\bar{D}^\kappa: \kappa \in H(\beta, l, r, L)\}$  lies within some Hölder class  $H(\beta, l', r', L')$ , with  $l', r', L'$  independent of  $\kappa$  (cf. lemma 7). This finishes the proof of theorem 1(a).

4.3. Proof of theorem 2

This subsection is devoted to the proof of theorem 2. The first step is to switch from the “stationary” model under consideration to the corresponding model with fixed initial point  $\Delta$ . As an immediate consequence of (37) in theorem 4 it is sufficient to prove (14)

$$\liminf_{t \rightarrow \infty} \inf_{K_t} R_t(K_t, H(\beta, l, r, L), t^{\beta/(2\beta+1)}) > 0$$

for the corresponding model with starting point  $\Delta$  where the risk is

$$R_t(K_t, H(\beta, l, r, L), r_t) = \sup_{\kappa \in H(\beta, l, r, L)} E_{\Delta}^{\kappa} \left( r_t^2 \int_I |K_t(a) - \kappa(a)|^2 da | \mathcal{A}_t(v, I) \right).$$

After this, our way towards theorem 2 follows a classical route, and we take advantage of the known structure of likelihoods in the model with fixed initial point; we even know about local asymptotic normality (LAN) available for this model (see Korostelev & Tsybakov, 1993; Hoffmann, 1999, for an approach of this type, and Löcherbach, 1999a, b, 2002, for the structure of likelihood ratios and LAN).

For simplicity, take  $I = [0, 1]$  (this is no loss of generality since (14) is not affected by affine transformations of space). Fix  $\beta, l, r, L$  as above, take  $r > 3l$ . Choose a function  $\chi$  which is  $\mathcal{C}^{\infty}$  on  $\mathbb{R}$ , non-negative, with compact support contained in  $[0, 1]$ . We consider finite families of type

$$\mathcal{C}_{j,\gamma} := \left\{ \kappa : \kappa = 2l + \gamma \sum_{k=0}^{2^j-1} \varepsilon_k \chi_{j,k}, \varepsilon_k \in \{0, 1\}, k = 0, 1, \dots, 2^j - 1 \right\} \tag{67}$$

where  $\chi_{j,k}$  is obtained from  $\chi$  by

$$\chi_{j,k}(y) = 2^{j/2} \chi(2^j y - k), \quad k = 0, 1, \dots, 2^j - 1, \quad y \in \mathbb{R}.$$

Note that functions  $\chi_{j,k}$  are supported by  $I_{j,k} := [k2^{-j}, (k+1)2^{-j}]$  and that  $\bigcup_{k=0}^{2^j-1} I_{j,k} = I$ . For  $t$  large enough, let the indices  $j, \gamma$  depend on  $t$  and take

$$j_t := \left\lfloor \frac{1}{2\beta+1} \log_2(t) \right\rfloor, \quad \gamma_t := t^{-1/2}. \tag{68}$$

Then  $2^{j_t}$  is of the order  $t^{1/(2\beta+1)}$  as  $t \rightarrow \infty$ . For  $t$  large enough we have

$$\mathcal{C}_{j_t, \gamma_t} \subset H(\beta, l, r, L).$$

Note also that  $|\mathcal{C}_{j_t, \gamma_t}| = 2^{2^{j_t}}$ . For  $k = 0, 1, \dots, 2^j - 1$ , denote by  $\mathcal{V}_t$  the space of sequences  $(\varepsilon_k)_{0 \leq k < 2^j}$  with entries  $\pm 1$ ,  $\mathcal{V}_{t,k}^0$  the space of such sequences with  $k$ th component deleted. Let  $\kappa_{u,k,+}$  be the function in  $\mathcal{C}_{j_t, \gamma_t}$  defined by the sequence  $u \in \mathcal{V}_{t,k}^0$  and by  $\varepsilon_k = +$ ;  $\kappa_{u,k,-}$  analogously with  $\varepsilon_k = -$ . Hence  $\kappa_{k,u,+} - \kappa_{k,u,-} = 2\gamma_t \chi_{j_t,k}$ . We write

$$L_t^{\kappa_{u,k,+}/\kappa_{u,k,-}} = \frac{d(Q_{\Delta}^{\kappa_{k,u,+}} | \mathcal{F}_t)}{d(Q_{\Delta}^{\kappa_{k,u,-}} | \mathcal{F}_t)}.$$

**Lemma 11**

Assume assumptions 4(l,  $\bar{p}$ ), 5(l), 6(l) and 7(1). Then the following condition (69) implies (14):

$$\limsup_{t \rightarrow \infty} \sup_{0 \leq k < 2^{j_t}} \sup_{u \in \mathcal{V}_{t,k}^0} E_{\Delta}^{\kappa_{k,u,-}} \left\{ \left| \log L_t^{\kappa_{k,u,+}/\kappa_{k,u,-}} \right| \right\} < \infty. \tag{69}$$

**Lemma 12**

Under assumptions 1–3, we have (69).

Combining lemmata 11 and 12, we have proved theorem 2.

**5. Concluding remarks**

Kernel estimation in non-parametric statistics has been extensively studied over the last 30 years. The book by Ibragimov & Khasminskii (1981) gives a good account of the state of the art around 1980 for models like density estimation under i.i.d. observations or Gaussian white noise. Our results show that the rate of convergence we obtain for branching diffusions are the same as in those classical models.

Recent directions of research in non-parametrics have mainly been improvement of estimation procedures through the notion of adaptivity (see e.g. the book of Härdle *et al.*, 1998, and the references therein). Our approach is non-adaptive in spirit: the construction of the optimal bandwidth  $t^{-1/(2\beta+1)}$  requires knowledge of the smoothness  $\beta$  of the unknown branching rate. Presumably, a general adaptive method (for instance the construction of Lepski, 1991) could be applied in our framework, at the cost of a still higher level of technicality, and inflating the rate of convergence by a log factor.

From a probabilistic point of view, the paper is purely one-dimensional. The configuration local time is an essential tool: it provides a Lebesgue density for the occupation time measure and ensures the continuity of the density of the invariant measure. An extension to higher dimensions requires a more sophisticated approach (see Bally & Löcherbach, 2001, for the study of the regularity of the invariant density in the multi-dimensional case).

**Acknowledgements**

We are grateful to two anonymous referees whose suggestions helped to improve a former version of the manuscript.

**References**

- Adke, S. R. (1983). Maximum likelihood estimation for a branching diffusion process. *J. Soviet Math.* **21**, 1–10.
- Athreya, K. B. & Ney, P. E. 1972. *Branching processes*. Springer-Verlag, Berlin, Heidelberg, New York.
- Bally, V. & Löcherbach, E. (2001). On the invariant density of branching diffusions. Preprint 688, Laboratoire de Probabilités, Université Paris VI.
- Crow, J. & Kimura, M. (1970). *An introduction to population genetics*. Harper & Row, New York.
- Daley, D. & Vere-Jones, E. (1988). *An introduction to the theory of point processes*. Springer Verlag, New York.
- Etheridge, A. M. (1993). Asymptotic behaviour of measure-valued critical branching processes. *Proc. Amer. Math. Soc.* **118**, 1251–1261.
- Ewens, W. (1969). *Population genetics*. Methuen, New York.
- Fleming, W. & Su, C. H. (1974). Some one-dimensional migration models in population genetics theory. *Theoret. Population Biol.* **5**, 431–449.
- Guttorp, P. (1991). *Statistical inference for branching processes*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York.
- Härdle, W., Kerkycharian, G., Picard, D. & Tsybakov, A. (1998). *Wavelets, approximation, and statistical applications*. Lecture Notes in Statistics, **129**, Springer Verlag, New York.
- Holmes, E. E. & Lewis, M. A. (1994). Partial differential equations in ecology: spatial interactions and populations dynamics. *Ecology* **75**, 17–29.
- Hoffmann, M. (1999).  $L_p$  estimation of the diffusion coefficient. *Bernoulli* **5**, 447–481.

- Höpfner, R. & Löcherbach, E. (1998). Birth and death on a flow: local time and estimation of a position-dependent death rate. *Stat. Inference Stoch. Process.* **1**, 225–243.
- Höpfner, R. & Löcherbach, E. (1999). On invariant measure for branching diffusions. Unpublished note, see <http://www.mathematik.uni-mainz.de/~hoepfner>.
- Höpfner, R. & Löcherbach, E. (2000). Limit theorems for null recurrent Markov processes. To appear in *Memoirs AMS*. Preprint 22-2000 at Mainz, Fachbereich Mathematik, University of Mainz.
- Ibragimov, I. A. & Khas'minskii, R. Z. (1981). *Statistical estimation— asymptotic theory*. Springer Verlag, Berlin.
- Jacod, J. (2000). Non-parametric kernel estimation of the diffusion coefficient of a diffusion. *Scand. J. Statist.* **27**, 83–96.
- Jacod, J. & Shiryaev, A.N. (1987). *Limit theorems for stochastic processes*. Springer, Berlin.
- Karatzas, J. & Shreve, S. (1991). *Brownian motion and stochastic calculus*, 2nd edn. Springer Verlag, New York.
- Khas'minskii, R. (1980). *Stochastic stability of differential equations*. Sijthoff & Noordhoff, Alphen.
- Kimura, M. & Weiss, G. (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **49**, 561–576.
- Korostelev, A. P. & Tsybakov, A. D. (1993). *Minimax theory of image reconstruction*. Lecture Notes in Statistics, **82**, Springer Verlag, New York.
- Kulperger, R. (1986). Parametric estimation for simple branching diffusion processes, II. *J. Multivariate Anal.* **18**, 225–241.
- Kutoyants, Y. (1998). *Statistical inference for spatial Poisson processes*. Lecture Notes in Statistics, **134**, Springer Verlag, New York.
- Lepski, O. V. (1991). Asymptotically minimax adaptive estimation I: upper bounds. Optimally adaptive estimates. *Theory Probab. Appl.* **36**, 682–697.
- Löcherbach, E. (1999a). Statistical models and likelihood ratio processes for interacting particle systems with branching and immigration. Thesis, Universität Paderborn.
- Löcherbach, E. (1999b). Likelihood ratio processes for Markovian particle systems with killing and jumps. Preprint 551, Laboratoire de Probabilités, Université Paris VI, to appear in *Stat. Inference Stoch. Process.*
- Löcherbach, E. (2002). LAN and LAMN for systems of interacting diffusions with branching and immigration. *Ann. Inst. H. Poincaré* **38**, 59–90.
- Malécot, G. (1967). Identical loci and relationship. *Proc. 5th Berkeley Symp. Math. Statist. Probab.* **4**, 317–332.
- Nagylaki, T. (1978). A diffusion model for geographically structured populations. *J. Math. Biol.* **6**, 375–382.
- Neyman, J. & Scott, E. (1964). A stochastic model of epidemics. In *Stochastic Models Med. Biol., Proc. Sympos. Univ. Wisconsin 1963* (ed. J. Gurland), pp. 45–83, University of Wisconsin Press, Madison.
- Pakes, A. (1975). On Markov branching processes with immigration. *Sankhya* **37**, 129–138.
- Revuz, D. & Yor, M. (1991). *Continuous martingales and Brownian motion*. Springer Verlag, Berlin.
- Sawyer, S. (1976). Branching diffusion processes in population genetics. *Adv. Appl. Probab.* **8**, 659–689.
- Sewastjanow, B. A. (1975). *Verzweigungsprozesse*. Oldenburg Verlag, München Wien.
- Wakolbinger, A. (1995). Limits of spatial branching processes. *Bernoulli* **1**, 171–189.
- Zubkov, A. (1972). Life-periods of branching processes with immigration. *Theory Probab. Appl.* **17**, 174–183.

Received April 2000, in final form October 2001

Eva Löcherbach, Centre de Mathématiques Appliquées, EA 2343, Université Paris XII—Val de Marne, UFR Sciences et Technologie, 61 avenue du Général de Gaulle, 94010 Créteil Cedex, France.  
E-mail: locherbach@univ-paris12.fr

## 6. Appendix

*Proof of Lemma 8.* (1) First, theorem 4, (36), implies that  $\inf_{\kappa \in H(\beta, l, r, L)} \inf_{a \in I} \bar{D}^\kappa(a) > 0$ . We prove for  $\varepsilon > 0$  arbitrary

$$\lim_t \sup_{\kappa \in H(\beta, l, r, L)} Q_m^\kappa \left\{ \sup_{a \in I} |A^{t, h_t}(a) - \bar{D}^\kappa(a)| > \varepsilon \right\} = 0. \quad (70)$$



By continuity of  $a \mapsto A^{t,h_t}(a)$  and  $a \mapsto \bar{D}^\kappa(a)$ ,

$$\begin{aligned}
 E_m^\kappa \left\{ \sup_{a \in I} (A^{t,h_t}(a) - \bar{D}^\kappa(a))^2 \right\} &= E_m^\kappa \left\{ (A^{t,h_t}(a_t(\omega)) - \bar{D}^\kappa(a_t(\omega)))^2 \right\} \\
 &\leq 2E_m^\kappa \left\{ \left( \int \frac{1}{h_t} Q \left( \frac{y - a_t(\omega)}{h_t} \right) \left( \frac{\eta_t}{t} - \bar{m}^\kappa \right) (dy) \right)^2 \right\} \\
 &\quad + 2E_m^\kappa \left\{ \left( \int \frac{1}{h_t} Q \left( \frac{y - a_t(\omega)}{h_t} \right) \bar{m}^\kappa (dy) - \bar{D}^\kappa(a_t(\omega)) \right)^2 \right\}.
 \end{aligned} \tag{71}$$

We consider the first term in this expression:

$$\begin{aligned}
 &E_m^\kappa \left\{ \left( \int \frac{1}{h_t} Q \left( \frac{y - a_t(\omega)}{h_t} \right) \left( \frac{\eta_t}{t} - \bar{m}^\kappa \right) (dy) \right)^2 \right\} \\
 &\leq E_m^\kappa \left\{ \int \frac{1}{h_t} Q \left( \frac{y - a_t(\omega)}{h_t} \right) \left( \frac{\bar{D}_t}{t} - \bar{D}^\kappa \right)^2 (y) dy \right\} \\
 &\leq \left[ \sup_{a \in K} \sup_{\kappa \in H(\beta, l, r, L)} E_m^\kappa \left\{ \left( \frac{\bar{D}_t}{t} - \bar{D}^\kappa \right)^2 (a) \right\} \right] \cdot \int_{\mathbb{R}} Q(x) dx
 \end{aligned}$$

for a suitable compact set  $K$ , since  $Q$  has compact support. Now corollary 1 yields the desired result. The second term in (71) is treated by using the fact that all  $\bar{D}^\kappa$  for  $\kappa \in H(\beta, l, r, L)$  are uniformly continuous (see lemma 7).

(2) Choose  $v < \inf_{\kappa \in H(\beta, l, r, L)} \inf_{a \in I} \bar{D}^\kappa(a)$  and  $\varepsilon > 0$  such that

$$v + \varepsilon < \inf_{\kappa \in H(\beta, l, r, L)} \inf_{a \in I} \bar{D}^\kappa(a).$$

Since  $\inf_{a \in I} A^{t,h_t}(a) \geq v$  on  $\{\sup_{a \in I} |A^{t,h_t}(a) - \bar{D}^\kappa(a)| \leq \varepsilon\}$  for all  $\kappa \in H(\beta, l, r, L)$ , we have by (70)

$$\begin{aligned}
 &\liminf_t \inf_{\kappa \in H(\beta, l, r, L)} Q_m^\kappa \left( \inf_{a \in I} A^{t,h_t}(a) \geq v \right) \\
 &\geq \liminf_t \inf_{\kappa \in H(\beta, l, r, L)} Q_m^\kappa \left( \sup_{a \in I} |A^{t,h_t}(a) - \bar{D}^\kappa(a)| \leq \varepsilon \right) = 1.
 \end{aligned}$$

*Proof of lemma 9.* It is sufficient to consider probability measures  $Q(u)du$  (otherwise decompose the signed finite measure  $Q(u)du$  into its positive and negative part and re-normalize to get probability measures  $(1/c^+)Q^+(u)du$ ,  $(1/c^-)Q^-(u)du$ ). Completing  $(\Omega, \mathcal{A}, \mathbb{F})$  with respect to  $Q_m$ , we have a process  $(\bar{D}(t, \cdot))_{t \geq 0}$  of Lebesgue densities for the process of occupation time measures  $(\eta_t)_{t \geq 0}$ . With Jensen inequality for the probability law

$$\frac{1}{h} Q \left( \frac{y - a}{h} \right) dy,$$

$f$  being continuous and  $Q$  with compact support, theorem 5 gives

$$\int_{\mathbb{R}} \frac{1}{h} Q \left( \frac{y - a}{h} \right) f(y) \left( \frac{\eta_t}{t} - \bar{m} \right) (dy) \rightarrow 0$$

in  $L^2(Q_m)$  as  $t \rightarrow \infty$  and  $h \rightarrow 0$ . The density  $\bar{D}$  of  $\bar{m}$  being continuous, the assertion is proved.

*Proof of lemma 10.* Consider  $(\mathbb{F}_{st})_{s \geq 0}$ -martingales  $(th)^{-1/2} Q((\cdot - a)/h) * (\mu^B - \nu^B)_{st}$  with angle bracket  $(th)^{-1} Q^2((\cdot - a)/h) * \nu_{st}^B$  and jumps bounded by  $cst(th)^{-1/2}$ . Using (7) and lemma 9, the martingale convergence theorem yields the assertion.

*Proof of lemma 11.* (1) For  $k$  fixed, we can write elements of  $\mathcal{V}_t$  in form  $(u, \pm)$  with  $u \in \mathcal{V}_{t,k}^0$ . Correspondingly, we write  $P_v$  for  $Q_A^k$  if  $v$  is the sequence of signs characterizing a branching rate  $\kappa = \kappa_v$  in  $\mathcal{C}_{j_i, \gamma_i}$ , or  $P_{u, \pm}$  if  $v$  is written as  $(u, \pm)$  with respect to some fixed  $k$ .

(2) Let  $K_t$  denote any  $\mathcal{F}_t$ -measurable estimator for  $\kappa$ . Obviously

$$\begin{aligned} R_t(K_t, H(\beta, l, r, L), t^{\beta/(2\beta+1)}) &\geq R_t(K_t, \mathcal{C}_{j_i, \gamma_i}, t^{\beta/(2\beta+1)}) \\ &\geq t^{2\beta/(2\beta+1)} \frac{1}{2^{2l}} \sum_{v \in \mathcal{V}_t} E_v \left( \int_I |K_t(a) - \kappa_v(a)|^2 da | \mathcal{A}_t(v, I) \right) \\ &= t^{2\beta/(2\beta+1)} \frac{1}{2^{2l}} \sum_{k=0}^{2l-1} \sum_{v \in \mathcal{V}_t} E_v \left( \int_{I_{j_i, k}} |K_t(a) - \kappa_v(a)|^2 da | \mathcal{A}_t(v, I) \right). \end{aligned}$$

Using the form of  $\kappa_v \in \mathcal{C}_{j_i, \gamma_i}$  determined by the sequence of signs  $v \in \mathcal{V}_t$ , we can rewrite the last expression as

$$\begin{aligned} t^{2\beta/(2\beta+1)} \frac{1}{2^{2l}} \sum_{k=0}^{2l-1} \sum_{u \in \mathcal{V}_{t,k}^0} (E_{u,+}(d_{t,k}^2(K_t, 2l - \gamma_t \chi_{j_i, k}) | \mathcal{A}_t(v, I)) \\ + E_{u,-}(d_{t,k}^2(K_t, 2l + \gamma_t \chi_{j_i, k}) | \mathcal{A}_t(v, I))) \end{aligned}$$

with notation  $d_{t,k}^2(f, g) = \int_{I_{j_i, k}} |f(a) - g(a)|^2 da$ .

(3) Now we look for a lower bound for the summands appearing in the last expression.

Define

$$\alpha_t := \frac{1}{2} d_{t,k}^2(2l - \gamma_t \chi_{j_i, k}, 2l + \gamma_t \chi_{j_i, k}) = \frac{1}{2} \gamma_t^2 \int_{\mathbb{R}} |\chi(u)|^2 du. \tag{72}$$

By the triangular inequality, this choice implies the following inclusion

$$\{d^2(K_t, 2l - \gamma_t \chi_{j_i, k}) \leq \alpha_t\} \subset \{d^2(K_t, 2l + \gamma_t \chi_{j_i, k}) \geq \alpha_t\}. \tag{73}$$

Then we have successively for every fixed  $k$  and every  $u \in \mathcal{V}_{t,k}^0$

$$\begin{aligned} E_{u,+}(d_{t,k}^2(K_t, 2l - \gamma_t \chi_{j_i, k}) | \mathcal{A}_t(v, I)) + E_{u,-}(d_{t,k}^2(K_t, 2l + \gamma_t \chi_{j_i, k}) | \mathcal{A}_t(v, I)) \\ \geq \alpha_t (P_{u,+}(\{d_{t,k}^2(K_t, 2l - \gamma_t \chi_{j_i, k}) \geq \alpha_t\} \cap \mathcal{A}_t(v, I)) \\ + P_{u,-}(\{d_{t,k}^2(K_t, 2l + \gamma_t \chi_{j_i, k}) \geq \alpha_t\} \cap \mathcal{A}_t(v, I))) \\ = \alpha_t E_{u,-}((L_t^{K_{u,+}/K_{u,-}} 1_{\{d_{t,k}^2(K_t, 2l - \gamma_t \chi_{j_i, k}) \geq \alpha_t\}} + 1_{\{d_{t,k}^2(K_t, 2l + \gamma_t \chi_{j_i, k}) \geq \alpha_t\}}) 1_{\mathcal{A}_t(v, I)}) \\ \geq \alpha_t E_{u,-}((L_t^{K_{u,+}/K_{u,-}} 1_{\{d_{t,k}^2(K_t, 2l - \gamma_t \chi_{j_i, k}) \geq \alpha_t\}} + 1_{\{d_{t,k}^2(K_t, 2l - \gamma_t \chi_{j_i, k}) \leq \alpha_t\}}) 1_{\mathcal{A}_t(v, I)}) \end{aligned}$$

where we have used (73); but the last expression is bounded below by

$$\alpha_t \exp(-s) P_{u,-}(\{L_t^{K_{u,+}/K_{u,-}} \geq \exp(-s)\} \cap \mathcal{A}_t(v, I)) \tag{74}$$

for arbitrary  $s > 0$ .

(4) We deduce from condition (69) that for  $s \in (0, \infty)$  large enough, there is some  $\delta > 0$  such that

$$P_{u,-}(\{L_t^{K_{u,+}/K_{u,-}} \geq \exp(-s)\} \cap \mathcal{A}_t(v, I)) > \delta \tag{75}$$

uniformly in  $k = 0, 1, \dots, 2^j - 1$ ,  $u \in \mathcal{V}_{t,k}^0$ , and  $t$  sufficiently large. Indeed

$$P_{u,-}(\{L_t^{\kappa_{u,+}/\kappa_{u,-}} \geq \exp(-s)\} \cap \mathcal{A}_t(v, I)) \geq P_{u,-}(|\log L_t^{\kappa_{u,+}/\kappa_{u,-}}| \leq s) + c_t - 1 \geq c_t - \frac{1}{s} E_{u,-}(|\log L_t^{\kappa_{u,+}/\kappa_{u,-}}|)$$

where  $c_t$  is a lower bound for all  $P_{u,-}(\mathcal{A}_t(v, I))$  such that  $\liminf_t c_t = 1$  for  $v$  sufficiently small, which exists by lemma 8, and where we used Chebyshev's inequality. Due to condition (69), we obtain (75) by taking  $s$  sufficiently large.

(5) Now we combine steps (2)–(4) to conclude the proof of lemma 11. Note that the double summation at the end of step (2) is over  $2^j 2^{2^j-1}$  terms, so we get

$$R_t(K_t, H(\beta, l, r, L), t^{\beta/(2\beta+1)}) \geq cst t^{2\beta/(2\beta+1)} 2^{j_i} \alpha_t$$

for all  $t$  sufficiently large. By (72) and by choice of  $j_t$  and  $\gamma_t$ , we do have that  $\alpha_t t^{2\beta/(2\beta+1)} 2^{j_i}$  is of the order  $t^{2\beta/(2\beta+1)} \gamma_t^2 2^{j_i}$ , which is bounded. Lemma 11 is proved.

*Proof of lemma 12.* From Löcherbach (1999a, b) we know that for any pair  $\kappa_{k,u,+}$ ,  $\kappa_{k,u,-}$  in  $\mathcal{C}_{j_i, \gamma_t}$  such that  $\kappa_{k,u,+} - \kappa_{k,u,-} = 2\gamma_t \chi_{j_i,k}$ , the log-likelihood ratio  $\log L_t^{\kappa_{k,u,+}/\kappa_{k,u,-}}$  has the form

$$\int_0^t \int_{I_{j_i,k}} \left( \log \frac{2l + \gamma_t \chi_{j_i,k}(y)}{2l - \gamma_t \chi_{j_i,k}(y)} \right) \mu^B(ds, dy) - \int_{I_{j_i,k}} 2\gamma_t \chi_{j_i,k}(y) \eta_t(dy). \tag{76}$$

Consider first the martingale part w.r.to  $\mathcal{Q}_\Delta^{\kappa_{k,u,-}}$  in (76):

$$\int_0^t \int_{I_{j_i,k}} \left( \frac{2l + \gamma_t \chi_{j_i,k}(y)}{2l - \gamma_t \chi_{j_i,k}(y)} - 1 \right) (\mu^B - \nu^{B, \kappa_{k,u,-}})(ds, dy),$$

where  $\nu^{B, \kappa_{k,u,-}}$  is the  $\mathcal{Q}_x^{\kappa_{k,u,-}}$ -compensator of  $\mu^B$ . Its angle bracket is

$$\int_{I_{j_i,k}} \frac{(2\gamma_t \chi_{j_i,k}(y))^2}{2l - \gamma_t \chi_{j_i,k}(y)} \eta_t(dy),$$

which by definition of  $\gamma_t$  and  $\chi_{j_i,k}$  is smaller than

$$Kt^{-1} 2^{j_i} \eta_t(I_{j_i,k}) \tag{77}$$

for some constant  $K$  not depending on  $t$  and  $k = 0, 1, \dots, 2^j - 1$ . In the same way, (77) is an upper bound for

$$\int_0^t \int_{I_{j_i,k}} \left( \frac{2l + \gamma_t \chi_{j_i,k}(y)}{2l - \gamma_t \chi_{j_i,k}(y)} - 1 \right)^2 \mu^B(ds, dy).$$

Write  $r(v) = \log(1 + v) - v + v^2/2$ . In (76), it remains to consider

$$\int_0^t \int_{I_{j_i,k}} r \left( \frac{2\gamma_t \chi_{j_i,k}(y)}{2l - \gamma_t \chi_{j_i,k}(y)} \right) \mu^B(ds, dy)$$

or by Lenglart domination the compensator of this expression. From  $r(v) = o(v^2)$  as  $v \rightarrow 0$  this compensator vanishes uniformly in  $k = 0, 1, \dots, 2^j - 1$  as  $t \rightarrow \infty$  in comparison to (77). So it remains to show

$$\limsup_{t \rightarrow \infty} \sup_{k=0,1,\dots,2^j-1} \sup_{u \in \mathcal{V}_{t,k}^0} E_\Delta^{\kappa_{k,u,-}} \left\{ \frac{1}{t} 2^{j_i} \eta_t(I_{j_i,k}) \right\} < \infty \tag{78}$$

in order to prove (69). Using again (37) in theorem 4, we may replace  $E_A^{K_{k,u,-}}(\cdot)$  in (78) by  $E_m^{K_{k,-}}(\cdot)$  and thus prove

$$\limsup_{t \rightarrow \infty} \sup_{k=0,1,\dots,2^j-1} \sup_{u \in \mathcal{I}_{t,k}^0} 2^{j_t} \bar{m}^{K_{k,u,-}}(I_{j_t,k}) < \infty. \tag{79}$$

We deduce (79) from lemma 1. From there, we have

$$\bar{m}^{K_{k,u,-}} = \pi \left( \sum_{n=1}^{\infty} (U^{K_{k,u,-}} \kappa_{k,u,-} \varrho)^n \right) U^{K_{k,u,-}}$$

where  $U^{K_{k,u,-}}$  is the occupation time kernel for the diffusion  $d\xi_t = b(\xi_t)dt + \sigma(\xi_t)dW_t$  killed at rate  $\kappa_{k,u,-}$ . We have by assumption 2 a bound  $\bar{\varrho} < 1$  for the reproduction means  $\varrho$ . Choose  $0 < a' < 2l < A' < \infty$  such that  $\varrho' := (A'/a')\bar{\varrho} < 1$ : then, as  $t \rightarrow \infty$ , we will have

$$a' < \kappa_{k,u,-}(\cdot) < A' \text{ uniformly in } k = 0, 1, \dots, 2^j - 1 \text{ for } t \text{ large enough.}$$

In particular,  $U^{K_{k,u,-}}$  is then smaller than the  $a'$ -potential kernel of the diffusion  $\xi$ . Thus all measures  $\bar{m}^{K_{k,u,-}}(dy) \kappa_{k,u,-}(y)$ ,  $k = 0, 1, \dots, 2^j - 1$ , are smaller than the  $a'(1 - \varrho')$ -potential kernel of the diffusion  $\xi$  integrated with respect to  $\pi$  (details analogous to Höpfner & Löcherbach, 1999, (2.3)). This potential kernel admits smooth Lebesgue densities. Since the intervals  $I_{j_t,k}$  have length  $2^{-j_t}$ , (79) is proved. Since we have seen that (79) implies (69), the proof of lemma 12 is finished.